

VŠB - TECHNICKÁ UNIVERZITA OSTRAVA
FAKULTA ELEKTROTECHNIKY A INFORMATIKY
KATEDRA MĚŘICÍ A ŘÍDICÍ TECHNIKY

Klasifikace typologie metabolismu na základě analýzy dat z energometrických testů

Disertační práce

2011

Ing. Michal Prauzek
školitel: prof. Ing. Vilém Srovnal, CSc.

Prohlášení

Prohlašuji, že jsem tuto disertační práci vypracoval samostatně.

Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

V Ostravě dne 14. 7. 2011

Ing. Michal Prauzek

Poděkování

Za vedení disertační práce děkuji svému školiteli prof. Ing. Vilému Srovnalovi, CSc. Za odborné konzultace a měření dat děkuji prof. MUDr. Karlu Martiníkovi, DrSc. z metabolické poradny v Hradci Králové. Za technickou podporu při měření děkuji Ing. Tomáši Martiníkovi. Za odborné konzultace děkuji Ing. Tomáši Peterkovi a Ing. Marku Penhakerovi, Ph.D. Za spolupráci na tématu a databázovém řešení děkuji Bc. Jakubu Hlavicovi. Za podporu a trpělivost při psaní této práce děkuji své rodině, přítelkyni a kamarádům. Práce byla podpořena stipendiem statutárního města Ostrava.

Abstrakt

Práce se zabývá klasifikací typologie metabolismu na základě analýzy dat z energometrických testů. V práci jsou použity tři základní zdroje dat získané od pacientů, jedná se jednak o měření respiračního kvocientu a následně vypočtené hodnoty utilizace potravy. Dále jsou to data biochemických rozborů krve a analýzy insulínu. A v poslední řadě data z měření impedance tkáně a popis pacienta. Pacienti jsou rozdělení na dvě skupiny, analytickou a testovací skupinu. Data pro analýzu jsou zpracována analýzou hlavních komponent, redukována a upravena pomocí normalizačních algoritmů. Práce rovněž zahrnuje metodiku pro klasifikaci a stanovení klasifikačních parametrů. Po vytvoření klasifikačního prostoru je algoritmus ověřen pomocí testovací skupiny a zhodnocen testováním statistických hypotéz. Výsledkem práce je algoritmus pro automatický diagnostický systém, který na základě získaného vzoru určí zařazení pacienta do určité skupiny souhrnných vlastností.

Klíčová slova: Analýza hlavních komponent, Biochemické testy, Fuzzy klasifikace, Insulín, Metabolismus, Měření bioimpedance, Normalizace dat, Respirační kvocient, Statistická hypotéza

This work deals with the classification of metabolism typology based on analysis of the energometry tests. Three data sources from patient are used in this work. The sources mean data from respiratory quotient measurement and calculated indicators of food utilization. Further the data from biochemical analysis of blood and insulin test are also used in analysis. At last, the measurement of bio-impedance and patient description are used in the work. The patients are divided into two groups, analytical and test group. The analysis group is processed by principal component analysis and it is normalized. The work also includes methods for fuzzy classification and estimation of classification parameters. After the classification space estimation the algorithm is tested by the testing group and evaluated by statistical hypothesis testing. The aim of the thesis is design algorithm for an automatic diagnostic system which insert patient to the group of features based on the previous model.

Key words: Principal component analysis, Biochemical tests, Fuzzy classification, Insulin, Metabolism, Bio-impedance measurement, Data normalization, Respiratory quotient, Statistical hypothesis

Seznam symbolů a zkratek

CO_2	Oxid uhličitý
O_2	Kyslík
ALT	Alaninaminotransferáza
AST	Aspartátaminotransferáza
ATH	Aktivní tělesná hmotnost
ATP	Adenozintrifosfát
BIA	Body impedance analysis
BMI	Body mass index
BMR	Basal metabolic rate, bazální metabolický obrat
BSA	Body surface area, plocha těla
C-P1	C-peptid na lačno
C-P2	C-peptid hodinu od podání glukózy
C-P3	C-peptid dvě hodiny od podání glukózy
CHOL	Cholesterol
CTV	Celková tělesná voda
ECT	Extracelulární tekutina
EE	Energetický ekvivalent
ER	Entity-relationship
GLU0	Krevní cukr na lačno
GLU1	Krevní cukr hodinu od podání glukózy
GLU2	Krevní cukr dvě hodiny od podání glukózy
GMT	Gamaglutamyltransferáza
HDLc	Lipoprotein s vysokou hustotou

ICT	Intracelulární tekutina
INS1	Insulín na lačno
INS2	Insulín hodinu od podání glukózy
INS3	Insulín dvě hodiny od podání glukózy
LDLV	Lipoprotein s nízkou hustotou
MOČ	Močovina
npRQ	Respirační kvocient bez zohlednění oxidace proteinů
RQ	Respirační kvocient
TAG	Triacylglycerol

Obsah

1	Metabolismus a energometrické testy	2
1.1	Intermediární metabolismus	2
1.2	Energetický metabolismus	3
1.3	Termoregulace	3
1.4	Energometrie	4
1.5	Respirační kvocient	5
1.6	Způsoby získávání energie	5
2	Měření a zpracování dat	7
2.1	Metabolický monitor Deltatrac II	7
2.2	Biochemický rozbor krve	11
2.3	Body status - QuadScan 4000	12
2.4	Metodika vyšetření	13
2.5	Deltatrac II, QuadScan 4000 - importování dat	13
2.6	Krevní testy - importování dat	14
2.7	Návrh databáze	15
3	Shrnutí současného stavu	18
3.1	Stávající stav řešení – řešerše	18
3.2	Předběžná a referenční studie	20
3.3	Současný stav vyhodnocení metabolického vyšetření	21
4	Cíl práce	22
4.1	Obecné cíle práce	22
4.2	Aplikační cíle	22
5	Návrh a realizace řešení	24
5.1	Možné způsoby řešení	24
5.2	Navržené řešení	24
5.3	Obecný model řešení	25
6	Analýza dat	26
6.1	Výběr pacientů	26
6.2	Popis dat pro analýzu a testování	26
6.3	Výběr metody analýzy dat	27
6.4	Analýza hlavních komponent	28
6.5	Návrh testovacích postupů	30

6.6	Test A - DeltaTrac II a biochemie	32
6.7	Test B - QuadScan 4000 a biochemie	34
6.8	Test C - DeltaTrac II, QuadScan 4000 a biochemie	36
7	Klasifikace dat	40
7.1	Tradiční metody klasifikace	40
7.2	Fuzzy metody klasifikace	40
7.3	Normalizace dat	41
7.4	Návrh počtu skupin	42
7.5	Klasifikační algoritmus	43
8	Návrh vyhodnocovacího algoritmu	46
8.1	Normalizace nových pacientů	46
8.2	Koeficienty analýzy hlavních komponent	46
8.3	Zařazení do klasifikačního prostoru	47
9	Testování navrženého algoritmu	50
9.1	Koncepce testování	50
9.2	Stanovení hranice příslušnosti	50
9.3	Testování statistických hypotéz	51
9.4	Vyhodnocení testování	51
9.5	Medicínské zhodnocení shluků	54
10	Závěr	58
10.1	Zhodnocení práce	58
10.2	Rekapitulace postupu řešení	59
10.3	Praktické uplatnění práce	60
10.4	Doporučení pro další postup	60
	Literatura	61
	Seznam publikací a výsledků autora	65
A	Rozdělení skupin pro analýzu a testování	I
B	Analýza hlavních komponent	III
C	Stanovení počtu shluků	VI
D	Grafická evaluace algoritmu	XI
E	Statistická evaluace výsledků	XVIII

Seznam obrázků

2.1	Přístroj Deltatrac II	7
2.2	Měřicí řetězec přístroje Deltatrac II	8
2.3	Přístroj QuadScan 4000	13
2.4	Algoritmus načítání dat - Deltatrac II	14
2.5	Algoritmus načítání dat - biochemické testy	15
2.6	ER diagram databáze	17
5.1	Obecný model	25
5.2	Diagnostický model	25
6.1	Rozložení věku všech pacientů pro analýzu a testování	27
6.2	Rozložení věku mužů pro analýzu a testování	28
6.3	Rozložení věku žen pro analýzu a testování	29
6.4	Informační hodnota hlavních komponent v celkovém testu	30
6.5	Informační hodnota hlavních komponent v testu A	31
6.6	Informační hodnota hlavních komponent v redukovaném testu A	32
6.7	Informační hodnota hlavních komponent v testu B	34
6.8	Informační hodnota hlavních komponent v redukovaném testu B	36
6.9	Informační hodnota hlavních komponent v testu C	38
6.10	Informační hodnota hlavních komponent v redukovaném testu C	39
7.1	Klasifikace v testu A	44
7.2	Klasifikace v testu B	45
8.1	Algoritmus zařazení do klasifikačního prostoru	47
8.2	Zařazení do klasifikačního prostoru v testu A	48
8.3	Zařazení do klasifikačního prostoru v testu B	49
10.1	Navržený postup řešení práce	59
A.1	Krabicový diagram: celkem, analýza a testování všech	I
A.2	Krabicový diagram: celkem, analýza a testování žen	II
A.3	Krabicový diagram: celkem, analýza a testování mužů	II
C.1	Indexy počtu shluků PC a CE - test A	VI
C.2	Indexy počtu shluků SC, S a XB - test A	VII
C.3	Indexy počtu shluků DI a ADI - test A	VII
C.4	Indexy počtu shluků PC a CE - test B	VIII
C.5	Indexy počtu shluků SC, S a XB - test B	VIII

C.6	Indexy počtu shluků DI a ADI - test B	IX
C.7	Indexy počtu shluků PC a CE - test C	IX
C.8	Indexy počtu shluků SC, S a XB - test C	X
C.9	Indexy počtu shluků DI a ADI - test C	X
D.1	První 4 parametry klasifikace testu A	XI
D.2	Druhé 4 parametry klasifikace testu A	XII
D.3	Poslední 2 parametry klasifikace testu A	XII
D.4	První 4 parametry klasifikace testu B	XIII
D.5	Druhé 4 parametry klasifikace testu B	XIII
D.6	Třetí 4 parametry klasifikace testu B	XIV
D.7	Poslední 4 parametry klasifikace testu B	XIV
D.8	První 4 parametry klasifikace testu C	XV
D.9	Druhé 4 parametry klasifikace testu C	XV
D.10	Třetí 4 parametry klasifikace testu C	XVI
D.11	Čtvrté 4 parametry klasifikace testu C	XVI
D.12	Páté 4 parametry klasifikace testu C	XVII
D.13	Poslední 3 parametry klasifikace testu C	XVII

Seznam tabulek

1.1	Hodnoty respiračního kvocientu	5
2.1	Obsahy energií ve složkách potravy	9
2.2	Parametry biochemického rozboru	12
2.3	Parametry výstupu přístroje QuadScan 4000	14
2.4	Typologie vyšetření	15
2.5	Struktura databáze	16
6.1	Parametry celkového testu	29
6.2	Parametry testu A	32
6.3	Parametry redukovaného testu A	33
6.4	Informační hodnota hlavních komponent - test A	34
6.5	Parametry testu B	35
6.6	Parametry redukovaného testu B	35
6.7	Informační hodnota hlavních komponent - test B	35
6.8	Parametry testu C	37
6.9	Parametry redukovaného testu C	37
6.10	Informační hodnota hlavních komponent - test C	37
9.1	Evaluace navržených testů	53
9.2	Medicínské zhodnocení, shluk A	55
9.3	Medicínské zhodnocení, shluk B	56
9.4	Medicínské zhodnocení, shluk C	57
B.1	Redukce testu A	III
B.2	Redukce testu B	IV
B.3	Redukce testu C	V
E.1	Testování hypotéz – souhlasný test A	XVIII
E.2	Testování hypotéz – křížový test A	XIX
E.3	Statistické parametry shluků – test A	XIX
E.4	Testování hypotéz – souhlasný test B, část 1	XX
E.5	Testování hypotéz – souhlasný test B, část 2	XX
E.6	Testování hypotéz – křížový test B, část 1	XX
E.7	Testování hypotéz – křížový test B, část 2	XXI
E.8	Statistické parametry shluků – test B, část 1	XXI
E.9	Statistické parametry shluků – test B, část 2	XXII
E.10	Testování hypotéz – souhlasný test C, část 1	XXII

E.11 Testování hypotéz – souhlasný test C, část 2	XXIII
E.12 Testování hypotéz – souhlasný test C, část 3	XXIII
E.13 Testování hypotéz – křížový test C, část 1	XXIII
E.14 Testování hypotéz – křížový test C, část 2	XXIV
E.15 Testování hypotéz – křížový test C, část C	XXIV
E.16 Statistické parametry shluků – test C, část 1	XXV
E.17 Statistické parametry shluků – test C, část 2	XXV
E.18 Statistické parametry shluků – test C, část 3	XXVI

Úvod

Civilizační nemoci jsou v dnešní době velkým problémem pro celou lidskou společnost. Čím je společnost vyspělejší, tím civilizační nemoci způsobují komplikace ve větším měřítku. Nepředpokládá se, že budou tyto jevy v budoucnu ustupovat, naopak s rostoucím tempem životního stylu se budou objevovat stále nové nemoci.

Metabolické poruchy, obzvláště obezita, jsou typičtí zástupci civilizačních nemocí, které trápí většinu obyvatelstva ve vyspělých zemích světa. Lidské tělo není nastaveno na životní styl, který se vyvinul ve dvacátém století a stále se zrychluje. Samozřejmě někteří jedinci snášejí tento způsob života lépe, jiní však hůře. To se pak projevuje nejenom na obezitě jednotlivce, ale i na náchylnosti k nemocem a taky na délce života.

Moderní medicína nezvládne toto zvrátit, může však řádně diagnostikovat jednotlivce a doporučit správnou úpravu potravy a životního stylu, protože každý jedinec je individuální a reaguje různě na potravu a další podněty okolí. Výzkum v této oblasti může popsat typologii člověka v závislosti na typu metabolismu a pomoci stanovit doporučenou diagnózu pro jednotlivé skupiny. K vyšetření metabolismu pacienta se využívá energometrických testů, biochemických rozborů, měření impedance tkáně a dalších doplňkových metod, které jsou hodnoceny subjektivně lékařem bez standardizovaného postupu.

Tato práce se zabývá souhrnným zpracováním rozsáhlého souboru měření energometrických testů, biochemických rozborů, měření impedance a dalších dat, získaných měřeními v metabolické poradně. Veškerá pozorování jsou prováděna dle standardní metodiky, která funguje již od roku 2003.

Výsledek práce přispěje ke zlepšení diagnostiky pacientů a vytvoření standardního procesu hodnocení energometrických testů pomocí moderních matematických metod a nových výpočetních prostředků. V minulosti se práce zabývaly hodnocením energometrických testů z hlediska malé skupiny vybraných objektů. Tato práce se věnuje ucelené klasifikaci na rozsáhlém vzorku pacientů s velkým počtem měřených parametrů, včetně popisu metodiky měření, zpracování dat, návrhu klasifikačního algoritmu a ověřování výsledků.

Kapitola 1

Metabolismus a energometrické testy

Metabolismus je definován jako látková přeměna v živé hmotě, která zahrnuje chemicko-fyzikální děje probíhající v buňce, orgánu až celém organismu. Úroveň metabolismu se neustále mění, a to v závislosti jak na vnějších, tak i vnitřních podmínkách, či potřebách a stavu daného organismu [40].

1.1 Intermediární metabolismus

Intermediárním metabolismem je nazývána přeměna hlavních složek výživy na buněčné úrovni, která slouží k zabezpečení základních životních funkcí buňky. Jedná se o přeměnu [26]:

- sacharidů (např. glukóza),
- lipidů (glycerol a mastné kyseliny),
- proteinů (aminokyselin).

Metabolické dráhy jsou série chemických reakcí, které se vyskytují uvnitř buňky. Metabolické dráhy mají přesnou časovou a funkční hierarchii a jsou vzájemně podmíněny. Vztahy jednotlivých metabolických drah jsou velmi složité [36].

Zdrojem energie jsou živiny obsažené v potravě, které jsou enzymaticky rozkládány a vstřebávány v trávicí soustavě. Sacharidy se štěpí na jednoduché cukry (monosacharidy). Nejvýznamnější je glukóza. Tuhy (lipidy) jsou rozloženy na mastné kyseliny a glycerol, bílkoviny (proteiny) na aminokyseliny. Tyto jednoduché látky pak vstupují do složitých transformačních procesů intermediárního metabolismu, kde je energie vázaná v různých součástech potravy měněna v jeden společný využitelný zdroj - adenosintrifosfát - ATP [22].

Na intermediární metabolismus účinkuje velké množství hormonů, cytokinů a jiných biologických substancí, které různým způsobem modifikují a regulují jednotlivé biochemické kroky, které pak ve vzájemném účinku harmonizují dílčí i konečné metabolické produkty. Jako ukázkový případ se uvádí působení insulínu. Ten působí jak na metabolismus glycidů, tak na bílkoviny a tuhy. Zvýšená hladina insulínu za fyziologických podmínek snižuje hladinu glukózy. Insulín je též důležitý pro syntézu bílkovin a brání lipolýze (uvolňování mastných kyselin a glycerolu z triacylglycerolů) z tukové tkáně [26].

Na intermediární metabolismus významným způsobem působí stresové hormony, které umožňují adaptaci a přežití organismu. Jsou to katecholaminy (adrenalin a noradrenalin), které ovlivňují intermediární metabolismus snižováním sekrece insulínu v beta-buňkách Langerhansových ostrůvků pankreatu, pravděpodobně prostřednictvím adrenergních alfa-receptorů. Pro metabolické pochody je důležitý i druhý hormon pankreatu, a to glukagon. Působí opačně než insulín; uvolňuje glukózu ze zásobní energie glykogenu, a to cestou glykogenolýzy. Růstový hormon pak navíc zvyšuje hypoglykémii navozenou insulínem [26].

Vzájemná integrita a bezchybnost složitých regulačních mechanismů znamená pro praxi zdraví a dlouhověkost na straně jedné, na druhé straně dochází při selhávání jednotlivých metabolických drah se vznikem chybných regulací nejprve ke vzniku a manifestaci tzv. rizikových faktorů (zvýšení glykémie, poruchy tukového metabolismu atd.). V další etapě se porucha projevuje jako konkrétní klinické onemocnění s orgánovým postižením, například srdeční infarkt, nebo dokonce selhání srdce až smrt [26].

1.2 Energetický metabolismus

Energetický metabolismus se dělí z hlediska fyziologických reakcí na děje, kdy dochází k ukládání energie do rezerv (anabolické reakce), které vedou ke vzniku nových sloučenin, tkání cestou biosyntézy, a katabolické reakce, které vedou k degradaci buněk tkání [26].

Anabolismus je tvorba látek, při které se energie spotřebovává, nabídka substrátů je vyšší než jejich okamžitá potřeba. Vytvářejí se energetické rezervy, dochází k obnově a novotvorbě tkání. Anabolické děje převažují v situacích, kdy je tělesná aktivita omezena. Intenzitu metabolických dějů v organismu, charakterizovanou výdejem energie v klidu za přesně stanovených podmínek (ráno vleže před opuštěním lůžka, na lačno, při fyziologické teplotě těla a neutrální teplotě okolí), se označuje termínem bazální metabolismus. Jeho hodnota závisí na věku, pohlaví a velikosti těla [22].

Katabolismus je rozklad látek za současného uvolnění energie. Je charakterizován chyběním rezerv glykogenu a mobilizací nesacharidových zdrojů energie - tuků a bílkovin. Souvisí s vyšší aktivitou sympatiku. Probíhá při zvýšení tělesné pohybové aktivity a při udržování životních funkcí [22].

Vlastní energetický metabolismus je definován jako systém přeměn energie ze živin za vzniku energie. Charakteristická je přeměna makroergních fosfátových vazeb, která umožňuje průběh intrabuněčných metabolických pochodů [26].

1.3 Termoregulace

Život je možný jen v určitém teplotním rozmezí. Čím jsou organismy složitější, tím užší je pro ně rozmezí optimální teploty. Protože teplota je důležitým faktorem ovlivňující aktivitu bílkovin a tekutost membrán, vyšší organismy vynakládají úsilí ve snaze udržet teplotu svého vnitřního prostředí relativně stálou [25].

Termoregulace je schopnost organismu udržovat stálou optimální tělesnou teplotu. Na teplotě závisí veškeré biochemické pochody v organismu. Metabolické procesy se zrychlují nebo zpomalují podle toho, jestli se teplota zvyšuje nebo snižuje. Při urychlení metabolismu pak člověk hubne a naopak. Nastavení termogeneze je klíčový problém vzniku obezity spojený s beta 3 adrenergním receptorem [26].

Kůže, a zejména podkožní vrstvy s velkým obsahem tukové tkáně, představují účinnou izolaci bránící ztrátám tepla zejména vedením do okolního prostředí. Termoregulační řídicí systém má obvyklé prvky regulačního obvodu – termoreceptory, hypotalamická řídicí centra a termoregulační efektorové mechanismy [25].

Normální tělesná teplota lidského těla měřená v axile kolísá mezi 35,8 až 37,0 stupňů Celsia, v závislosti na aktivitě a stavu organismu, okolní teplotě, vlhkosti a proudění vzduchu v okolí a na oblečení měřeného. Této teplotě se říká teplota povrchu těla, to jsou ty části těla, jejichž tělesná teplota se částečně mění vlivem okolí (končetiny, hlava, povrchové části těla). Teplota jádra je relativně konstantní, nezávislá na teplotě okolí. Je to teplota v hrudní a břišní dutině. Její nejvyšší hodnota v játrech se pohybuje mezi 39-40 stupni Celsia. Teplota jádra je udržována ve stálém rozmezí díky tomu, že je od okolí izolována kůží, podkožním vazivem a tukovou vrstvou [26].

Tělesná teplota je udržována tvorbou a výdejem tepla. Jestliže je tvorba tepla vyšší než tepelné ztráty, teplota organismu se zvyšuje (vzniká horečka) a naopak. Klidová tvorba tepla probíhá hlavně v jádře těla, především v játrech. Většinou se teplo tvoří jako vedlejší produkt při metabolických dějích, může se však tvořit i cíleně: svalovou činností nebo zrychlením metabolismu účinkem metabolických hormonů. Výdej tepla je možný, pokud se teplo jádra odvádí cirkulující krví do kůže – povrchních struktur těla [26].

Za normálních okolností je tvorba a výdej tepla v rovnováze a udržuje se v rozsahu teplotní pohody. Termoregulace nastupuje až po přestoupení hranic teplotní pohody. Teplota se pak udržuje zpětnovazebním mechanismem, jehož centrum je v hypotalamu. Centrální tepelné receptory – detektory teploty, důležité pro zpětnovazební řízení, byly prokázány v hypotalamu samotném. Periferní tepelné receptory se nacházejí v kůži a v některých tkáních lidského těla [26].

1.4 Energometrie

Energie se v lidském těle vytváří cestou oxidace základních složek potravy nebo oxidací vlastních energetických substrátů. Při utilizaci jednotlivých složek stravy vznikají přesně definované energetické ekvivalenty a biochemické substráty. Lidské tělo přeměňuje asi 45 % energetického obsahu oxidovaných substrátů na práci a zbylých 55 % na teplo. Energii, která se přeměňuje na práci, je dále možno rozlišit na práci vnitřní (respirace, membránový transport, syntetické procesy atd.) a práci vnější (svalová práce při pohybu). Vzhledem k tomu, že i takzvaná vnitřní práce je v konečném důsledku přeměněna na teplo, je možno za klidových podmínek odhadovat potřebu energie pomocí energometrie [26].

Přímá energometrie je využívána jen na několika pracovištích na světě ve vědecké praxi, zatímco nepřímá energometrie je daleko používanější, protože sleduje produkty oxidačních procesů. V praxi se užívá měření metodou nepřímé kalorimetrie, kdy je spotřeba živin (výživových substrátů) vypočítána ze spotřeby kyslíku a výdeje oxidu uhličitého. Oxidace jednotlivých složek výživy je spojena se specifickými hodnotami spotřeby kyslíku a produkci oxidu uhličitého. Odpad dusíku do moči je pak ekvivalentní rychlosti degradace bílkovin. Z naměřených hodnot vypočítáváme klidový respirační kvocient, nazývaný také metabolický či utilizační ekvivalent, který je dán vztahem 1.1 [12],[32].

$$RQ = \frac{V_{CO_2}}{V_{O_2}} \quad (1.1)$$

1.5 Respirační kvocient

Respirační kvocient (RQ) je poměr mezi objemem vytvořeného CO_2 a objemem spotřebovaného O_2 za jednotku času při ustáleném stavu. Standardní hodnoty respiračního kvocientu pro jednotlivé složky potravy a celku jsou uvedeny v tabulce 1.1 [26], [42].

Tabulka 1.1: Hodnoty respiračního kvocientu

Složka potravy	Respirační kvocient
Tuky	0.70
Běžná potrava	0.80
Bílkoviny	0.82
Sacharidy	1.00
Kyselina jablečná	1.33
Kyselina vinná	1.60
Kyselina šťavelová	4.00

Určení respiračního kvocientu při spalování bílkovin v těle je složitější, jelikož při metabolismu bílkovin nevzniká pouze oxid uhličitý a voda, ale zároveň i jednoduché sloučeniny dusíku, které nejsou z těla vylučovány dýchacím ústrojím [42].

Respirační kvocient se liší v různých podmínkách, například při hyperventilaci (prohloubené dýchání) stoupá, protože je vydechován ve zvýšené míře oxid uhličitý. Při intenzivní námaze RQ stoupá, protože kyselina mléčná, která se produkuje při námaze, se mění na CO_2 , a ten je ve zvýšené míře vydechován. Při acidóze (snížení pH pod 7,36, zvýšení kyselosti krve) RQ stoupá a při alkalóze (zvýšení pH nad 7,44, snížení kyselosti krve) klesá [42].

Hodnoty RQ v jednotlivých orgánech významně charakterizují metabolické procesy, které se v nich odehrávají. Například RQ mozku je obvykle 0,97-0,99, což ukazuje, že jeho hlavním palivem, i když ne jediným, jsou sacharidy. Při sekreci žaludeční šťávy má žaludek negativní RQ, protože se vychytává z arteriální krve více CO_2 , než se odevzdává do krve venózní [42].

Při oxidaci sacharidů vzniká z každé spotřebované molekuly kyslíku jedna molekula oxidu uhličitého. Poměr spotřebovaného kyslíku a vytvořeného oxidu uhličitého se tedy rovná 1. Při oxidaci tuků vzniká jen 7 molekul oxidu uhličitého na 10 molekul spotřebovaného kyslíku, respirační kvocient je tedy 0,7. Podobnou hodnotu má i průměrný respirační kvocient bílkovin, kde při spotřebě 10 molekul kyslíku vzniká zhruba 8 molekul oxidu uhličitého [25].

1.6 Způsoby získávání energie

Anaerobní způsob získávání energie je charakterizován možností svalových buněk vykonávat mechanickou práci při využívání energie uvolněné bez účasti kyslíku. Anaerobní zdroje energie využívá organismus v situacích, kdy není schopen zabezpečit dostatek energie efektivnějším aerobním způsobem. Aerobní způsob získávání energie je dominantní při tělesných aktivitách vytrvalostního charakteru trvajícího déle než 23 minuty. Úroveň aerobních schopností je ovlivněna dědičností (80%) [22].

Anaerobní práh je taková intenzita zátěže (rychlost běhu, výkon na ergometru), kdy se k převážně aerobnímu způsobu krytí energetických požadavků organismu přidává a dále na-

růstá způsob anaerobní. Důsledkem je zvýšená tvorba a kumulace kyseliny mléčné se zvýšením koncentrace vodíkových iontů. Stálá hladina vodíkových iontů je nezbytnou podmínkou, nutnou k funkci buněk organismu. Proto dochází k aktivaci kompenzačních mechanismů, které udržují stálost vnitřního prostředí (homeostázu). Nejdůležitější je nárazníkový bikarbonátový systém. Slabá kyselina uhličitá disociuje na vodíkový iont a hydrouhličitan sodný (bikarbonát), současně je však v dynamické rovnováze s rozpuštěným oxidem uhličitým [22].

Kapitola 2

Měření a zpracování dat

2.1 Metabolický monitor Deltatrac II

Metabolický monitor Deltatrac II je nepřímý kalorimetr, který je konstruován pro klinické použití a zdravotnický výzkum. Systém je možno použít pro terapii výživy, intenzivní péči, diagnostiku diabetu a výzkum v oblasti metabolismu a chorobné obezity.



Obrázek 2.1: Přístroj Deltatrac II

Monitor měří metabolické parametry při umělé i spontální plicní ventilaci. Rozličné oblasti provozu umožňují monitorování pacientů od novorozenců až po dospělé pacienty s extrémní nadváhou. Základními parametry přístroje jsou spotřeba kyslíku, vylučování oxidu uhličitého, respirační kvocient, energetický obrat a inspirační koncentrace kyslíku. Přístroj Deltatrac II je zobrazen na obrázku 2.1 [13].

Přístroj používá infračervený analyzátor oxidu uhličitého a rychlý diferenční paramagnetický analyzátor kyslíku. Měření probíhá na základě analýzy při konstantním průtoku [27].

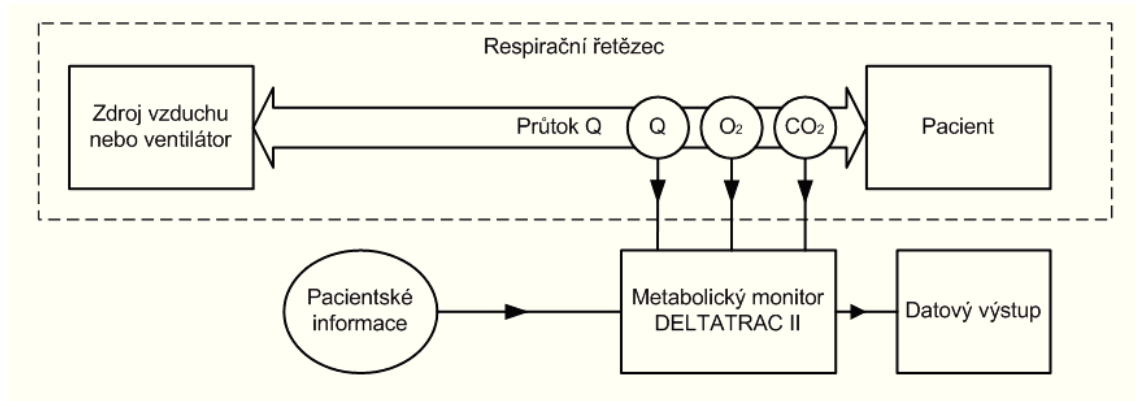
2.1.1 Kalibrace přístroje

Metabolický monitor je doporučeno kalibrovat jednou denně. Pro dosažení optimální přesnosti by se měl kalibrovat před každým měřením. Před vlastní kalibrací je předepsaná čekací doba pro zahřátí přístroje v délce 30 minut [13].

Při kalibraci je automaticky prověřována základní úroveň pro senzor kyslíku a senzor oxidu uhličitého. Metabolický monitor Deltatrac II má možnost automatické kalibrace s využitím interního hadicového propojení na magnetický ventil spojený s plynovou kalibrační láhví [13].

2.1.2 Teoretické podklady měření

Při měření za spontánního dýchání se hlava zakryje průhlednou plastovou maskou. Přístroj Deltatrac II vyrábí konstantní průtok 40 litrů vzduchu za minutu, proudícího skrze masku a přístroj. Z konstantního průtoku je vypočítána spotřeba kyslíku a produkce oxidu uhličitého, řetězec je zjednodušeně zobrazen v obrázku 2.2.



Obrázek 2.2: Měřicí řetězec přístroje Deltatrac II

Standardní respirační kvocient je vypočítán dle vzorce 1.1. Respirační kvocient bez zohlednění oxidace proteinů se vypočítává dle vzorce 2.1, kde V_{CO_2} je produkce oxidu uhličitého v ml/min, V_{O_2} spotřeba kyslíku v ml/min a U_N vyloučený dusík v moči v g/24h [13].

$$npRQ = \frac{1,44.V_{CO_2} - 4,89.U_N}{1,44.V_{O_2} - 6,04.U_N} \quad (2.1)$$

Klidový energetický obrat (EE) v kcal/24h je vypočítán dle vzorce 2.2, kde V_{CO_2} je produkce oxidu uhličitého v ml/min, V_{O_2} spotřeba kyslíku v ml/min a U_N vyloučený dusík v moči v g/24h. Stejným způsobem lze definovat i jednotlivé složky: sacharidy (vzorec 2.3), tuky (vzorec 2.4) a bílkoviny (vzorec 2.5) [13].

$$EE = 5,5.V_{O_2} + 1,76.V_{CO_2} - 1,99.U_N \quad (2.2)$$

$$CRO = 5,926.V_{CO_2} - 4,189.V_{O_2} - 2,539.U_N \quad (2.3)$$

$$FAT = 2,432.V_{O_2} - 2,432.V_{CO_2} - 1,943.U_N \quad (2.4)$$

$$PRO = 6,25.U_N \quad (2.5)$$

Pro výpočet proporcionálních podílů základních tří nosičů energie jsou použity obsahy energií uvedené v tabulce 2.1 [13].

Tabulka 2.1: Obsahy energií ve složkách potravy

Nosič energie	Energetická hodnota (kcal/g)
Tuky	9,46
Bílkoviny (Protein)	4,32
Sacharidy (Hydrát uhlíku)	4,18

BMR je energetický obrat definovaný pro zdravou osobu v klidu minimálně 12 hodin, při nedostatku potravin a normální okolní teplotě. Základní metabolický obrat (BMR v kcal/24h) lze vypočítat pomocí vztahu Harris-Benedict zvlášť pro muže (vzorec 2.6) a ženy (vzorec 2.7), kde W je váha v kg, H je výška v cm a A je stáří v letech [13].

$$BMR_{muži} = 66 + 13,8.W + 5.H - 6,8.A \quad (2.6)$$

$$BMR_{ženy} = 655 + 9,6.W + 1,8.H - 4,7.A \quad (2.7)$$

Rovnici Harris-Benedict lze použít pouze pro osoby starší 17 let. Další rovnice navržená Fleischem platí pro pacienty od věku 1 roku. Každá z rovnic je navržena pro různý věk osoby a pohlaví, BMR v kcal/24 lze vypočítat pro muže ve věku 1-12 let (vzorec 2.8), 13-19 let (vzorec 2.9) a 20-99let (vzorec 2.10). Dále pro ženy ve věku 1-10 let (vzorec 2.11), 11-19 let (vzorec 2.12) a 20-99 let (vzorec 2.13), kde A je věk v letech a BSA je povrch těla v cm^2 [13].

$$BMR_{muži1-12} = 24.BSA.(54 - 0,885.A) \quad (2.8)$$

$$BMR_{muži13-19} = 24.BSA.(42,5 - 0,642.(A - 13)) \quad (2.9)$$

$$BMR_{muži20-99} = 24.BSA.(38 - 0,073.(A - 20)) \quad (2.10)$$

$$BMR_{muži1-10} = 24.BSA.(54 - 1,045.A) \quad (2.11)$$

$$BMR_{muži11-19} = 24.BSA.(42,5 - 0,778.(A - 11)) \quad (2.12)$$

$$BMR_{muži20-99} = 24.BSA.(35,5 - 0,064.(A - 20)) \quad (2.13)$$

Pro malé děti ve věku do jednoho roku, novorozence a nedonošence se používá rovnice dle Schoefielda, která zohledňuje pouze váhu (W v kg), výpočet je uveden ve vzorci 2.14 [13].

$$BMR = 54.W \quad (2.14)$$

Tělesný povrch (BSA v cm^2) je u dospělých vypočítáván vztahem dle DuBois (vzorec 2.15) a u dětí vztahem dle Haycocka (vzorec 2.16). Ve vzorci je zohledněna váha (W) v kg, výška (H) v cm a použity konstanty C_D s hodnotou 71,84 a C_H s hodnotou 242,65 [13].

$$BSA = W^{0,425} \cdot H^{0,725} \cdot C_D \quad (2.15)$$

$$BSA = W^{0,5378} \cdot H^{0,3964} \cdot C_H \quad (2.16)$$

2.1.3 Datový výstup

Datový výstup má formát textového souboru, kde jsou v jednotlivých řádcích uvedeny údaje o měření. Hlavička dokumentu obsahuje informace o pohlaví pacienta, datum narození, věk, váhu, výšku a indexy BSA, BMR. Dále jsou součástí souboru informace o nastavení metabolického monitoru, okolní podmínky, čas měření a kalibrační výsledky. Část s výsledky obsahuje změřené a statisticky zpracované údaje o spotřebovaném kyslíku a produkováném oxidu uhličitým. Z těchto výsledků jsou potom počítány indexy respiračního kvocientu a energetického obratu rozděleného na složky (carbohydrate, fat, protein).

Ukázka datového výstupu přístroje Deltatrac II ve formě textového souboru:

DELTATRAC Metabolic Monitor Prog. rev. 882123-2.2

PATIENT:

SEX	female
DATE OF BIRTH	07-Jun-59
AGE	48 yr
HEIGHT	170 cm
WEIGHT	99 kg
NITROGEN EXCRETION	13 g/24h
BODY SURFACE AREA	2.10 m2
BASAL METABOLIC RATE	1690 kcal/24h (H-B)
	1700 kcal/24h (Fleisch)

CANOPY MODE ADULT

RESULTS IN STPD

NO ARTIFACT SUPPRESSION

GAS TEMPERATURE 31.1 C

AMBIENT PRESSURE 749 mmHg

MEASUREMENT STARTED 09:01 26-Sep-2007

TILL 09:12 26-Sep-2007

DURATION 00:11

ARTIFACTS AND

INTERRUPTS 00:01

CALIBRATIONS:	AMB. CO2	CO2 MEAS./SET	O2 MEAS./SET
10-Sep-2007 05:57	0.04 %	5.01 / 5.00 %	94.9 / 95.0 %

```

FLOW SETTING:      ADULT    40.1 l/min
MEASUREMENT RESULTS:
                    VC02      V02      RQ      EE
                    ml/min    ml/min
                    24h
MEAN                241      299      0.81    2040
STANDARD            23      20      0.07    130
DEVIATION            9.5 %    6.8 %    8.9 %    6.6 %
(N = 10)
DIFFERENCE FROM BASAL METABOLIC RATE:  + 21 % (H-B)
NON-PROTEIN RQ      0.80
ENERGY SUBSTRATE UTILIZATIONS:
                    g/24h      % OF TOTAL
CARBOHYDRATE        142      29.1
FAT                  116      53.7
PROTEIN              81      17.1
End of Report

```

2.1.4 Postup měření

Pacient přichází na měření odpočínut a na lačno. Vyšetření probíhá v několika fázích:

1. Uložení pacienta a uklidnění
2. Kalibrace
3. Samotné měření

Nejprve je pacient uložen na lůžko a je vyčkáno deset minut, aby se pacient uklidnil a adaptoval na prostředí. Poté je na pacienta nasazena respirační maska, která zahalí hlavu a horní část těla. Po upevnění masky je zvolen mód přístroje a provedena kalibrace. Pak následuje samotné měření, které má délku deset minut. Výsledkem měření jsou potom průměrné hodnoty respiračních plynů po celou dobu měření.

2.2 Biochemický rozbor krve

Biochemické krevní testy jsou další metodou k diagnostice metabolických procesů v organismu. Postupem času probíhají v lidském těle metabolické děje, které ovlivňují různé činitele v lidském organismu, které můžeme sledovat.

Nejčastěji se monitoruje hladina insulinu a C-peptidu po podání glukózy v pravidelných časových intervalech. Lze však sledovat i mnoho dalších ukazatelů v krevním rozboru. Velice dobrým ukazatelem metabolismu jsou i jaterní faktory ALT a AST, které určují míru funkčnosti a poškození jater, dále například přímá hodnota glukózy v krvi. V diagnostice používáme i další parametry jako jsou GMT, MOČ, TAG, LDLV, HDLC a CHOL. Součástí je i monitorování krevního cukru. Kompletní přehled parametrů krevního rozboru je uveden v tabulce 2.2 [24].

Tabulka 2.2: Parametry biochemického rozboru

Parametr	Jednotka
alt	$\mu\text{kat/l}$
ast	$\mu\text{kat/l}$
gmt	$\mu\text{kat/l}$
hdlc	mmol/l
ldlv	mmol/l
chol	mmol/l
moč	$\mu\text{mol/l}$
tag	mmol/l
glu0	mmol/l
glu1	mmol/l
glu2	mmol/l
ins1	pmol/l
ins2	pmol/l
ins3	pmol/l
c-p1	nmol/l
c-p2	nmol/l
c-p3	nmol/l

2.3 Body status - QuadScan 4000

Bodystat QuadScan 4000 je analyzátor složení těla, který využívá bioimpedanční měření lidského těla. Přístroj QuadScan 4000 je zobrazen na obrázku 2.3. Metoda funguje na principu měření obsahu vody v biologických strukturách celého organismu, a to konkrétně šířením elektrického proudu nízké intenzity lidským tělem. Tato diagnostika je neinvazivní a bezpečná pro pacienta [43].

Mezi základní popisné hodnoty patří věk, výška, váha a obvod pasu. Na základě stanovených frekvencí elektrického proudu QuadScan 4000 vypočítává parametry tělesného tuku, aktivní tělesné hmoty, celkové tělesné vody, intracelulárních a extracelulárních tekutin. Kompletní přehled parametrů přístroje QuadScan 4000 je uveden v tabulce 2.3 [43].

2.3.1 Postup měření

Na pacienta jsou připevněny čtyři elektrody k jeho pravé ruce a pravé noze. Pravá strana je vybrána z důvodu eliminace vlivu průchodu elektrického proudu srdcem. Elektrody jsou rozděleny do dvou skupin, A a B. V každé skupině se nachází černá a červená elektroda, jednotlivé skupiny jsou zaměnitelné. Červená elektroda se umísťuje na zánártí nebo zápěstí hned za konce prstů, černá se potom umísťuje blíže k srdci, a to na zápěstí a kotník [43].

Pacient se musí uvést do klidového stavu na dobu zhruba deseti minut před vyšetřením. Červené elektrody slouží jako zdroj proudu, černé pak jsou snímací. Jestliže nejsou správně připojeny, tak přístroj není schopen měřit a signalizuje toto na displeji. Měřený objekt nesmí být umístěn na vodivé podložce. Po řádném měření jsou údaje zobrazeny na displeji a lze je odeslat do softwaru v počítači [43].



Obrázek 2.3: Přístroj QuadScan 4000

2.3.2 Technologie měření

Měření impedance lidského těla (BIA) je jednou z metod klasifikace složení těla. Hmotnost lidského těla nevyjadřuje přesně obezitu a množství tuku v organismu. Pokud je nutné zjistit jednotlivé složky, pak je potřeba použít analýzu pomocí měření impedance [19].

K diagnostice se používá multifrekvenční analýza, kdy se postupně mění frekvence měření v několika krocích. Právě z těchto hodnot se potom vypočítávají jednotlivé parametry lidského organismu. Přístroj je napájen z baterie a proud je do těla dodáván pomocí elektrod. Přístroj měří impedanci pomocí frekvencí 5, 100 a 200 kHz [43].

2.4 Metodika vyšetření

Základní tři typy vyšetření pacientů, které se vyskytují v této práci, jsou uvedené v tabulce 2.4. Typ 1 je ukázkou kompletního vyšetření pacienta, kdy jsou provedeny všechny vstupní testy a po podání glukózy dva reakční testy v časovém období jedna a dvě hodiny. Vyšetření typu 2 je stejně strukturováno jako typ 1, ale je zjednodušeno o měření po jedné hodině. U vyšetření typu 3 se neprovádí test s podáním glukózy, pouze se testují hodnoty v normálním stavu.

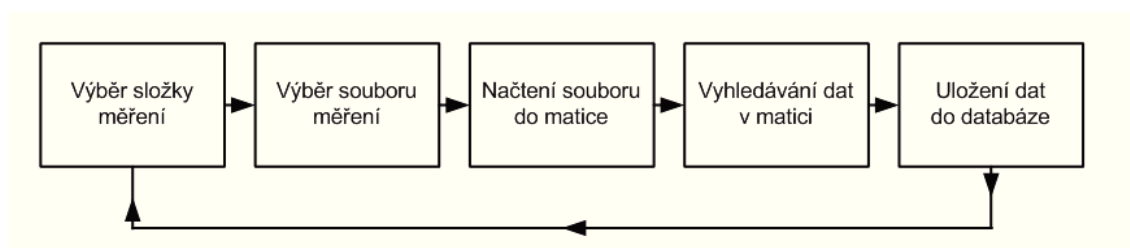
2.5 Deltatrac II, QuadScan 4000 - importování dat

Data generována měřicím systémem Deltatrac II jsou v textovém neformátovaném souboru, který musí být dále upraven. Vedle samotných dat jsou v souboru další doplňující údaje o pacientovi.

Textové informace je potřeba převést do číselného formátu. K tomu kroku byl zhotoven skript v prostředí MATLAB, který postupně načítá jednotlivé složky dnů měření. Z vybrané

Tabulka 2.3: Parametry výstupu přístroje QuadScan 4000

Parametr	Jednotka
věk	roky
výška	cm
hmotnost	kg
tukpc	%
tuk	kg
athpc	%
ath	kg
vodapc	%
voda	kg
bezvody	kg
ictpc	%
ectpc	%
ctvpc	%
ict	kg
ect	kg
ctv	kg



Obrázek 2.4: Algoritmus načítání dat - Deltatrac II

složky načítá jednotlivá měření, která splňují podmínky kompletnosti měření, a postupnou analýzou datové matice zjišťuje, ve kterém řádku se vyskytuje potřebná informace. Ze všech údajů sestavuje vektor měření, který je připraven k zápisu do databáze. Algoritmus zpracování souboru je zobrazen na obrázku 2.4.

Data ze systému QuadScan 4000 byla importována pomocí obdobného algoritmu jako u systému DeltaTrac II. Data bodystatu však byla v databázovém formátu, takže po importu dat do prostředí MATLAB mohlo ihned dojít k sestavení databázového příkazu a uložení do příslušné tabulky.

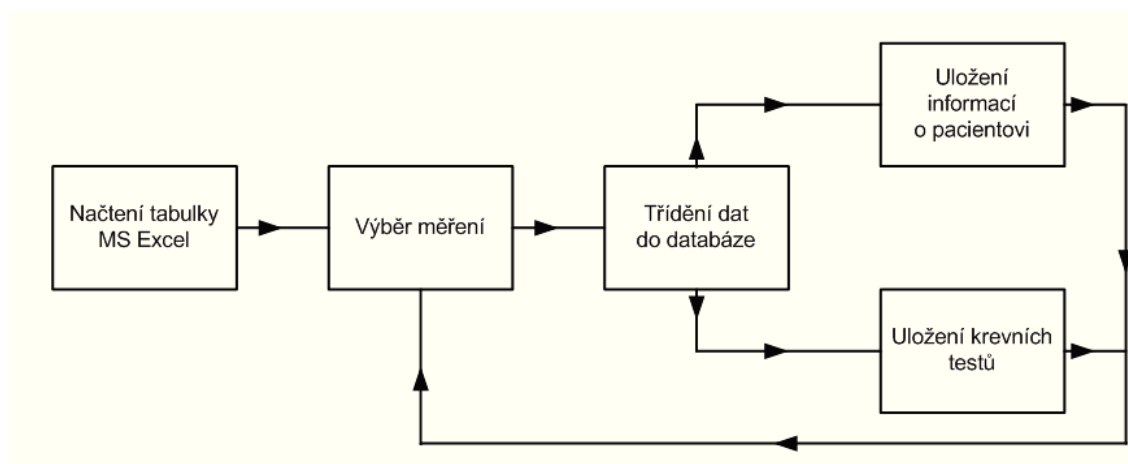
2.6 Krevní testy - importování dat

Data krevních testů jsou uložena v tabulce programu Microsoft Excel, kde je každému pacientovi přiřazen řádek s výsledky jednotlivých parametrů krevních testů v čase.

K uložení dat byla naprogramována procedura v prostředí MATLAB. V první fázi jsou

Tabulka 2.4: Typologie vyšetření

Prováděný úkon	Typ 1	Typ 2	Typ 3
Měření Deltatrac II - čas 0h	ANO	ANO	ANO
Měření QuadScan 4000 - čas 0h	ANO	ANO	ANO
Krevní obraz, insulín test 0h	ANO	ANO	ANO
Podání glukózy čas 0h	ANO	ANO	-
Měření Deltatrac II - čas 1h	ANO	-	-
Krevní obraz, insulín test 1h	ANO	-	-
Měření Deltatrac II - čas 2h	ANO	ANO	-
Krevní obraz, insulín test 2h	ANO	ANO	-



Obrázek 2.5: Algoritmus načítání dat - biochemické testy

data importována do prostředí MATLAB k následnému zpracování, ve druhém kroku je vybrán konkrétní pacient. Data jsou dále rozdělena do dvou bloků, jeden obsahuje informace o pacientovi, druhý pak konkrétní data měření. Ve finální fázi jsou data umístěna do jednotlivých tabulek databáze. Celý algoritmus zpracování souboru je zobrazen na obrázku 2.5.

K těmto datům byla přidána data krevního rozboru, která jsou uložena v tabulce biochemie a postupně byla rozdělena do dílčích parametrických tabulek.

2.7 Návrh databáze

Jako centrální úložiště dat byla zvolena MySQL databáze. Databáze obsahuje 18 základních tabulek, kde v první se nachází data měření z přístroje Deltatrac II a v dalších dvou data z insulínových testů, další tabulky obsahují údaje z měření bodystatu a další parametry krevního rozboru. Struktura a počty položek v databázi jsou uvedeny v tabulce 2.5.

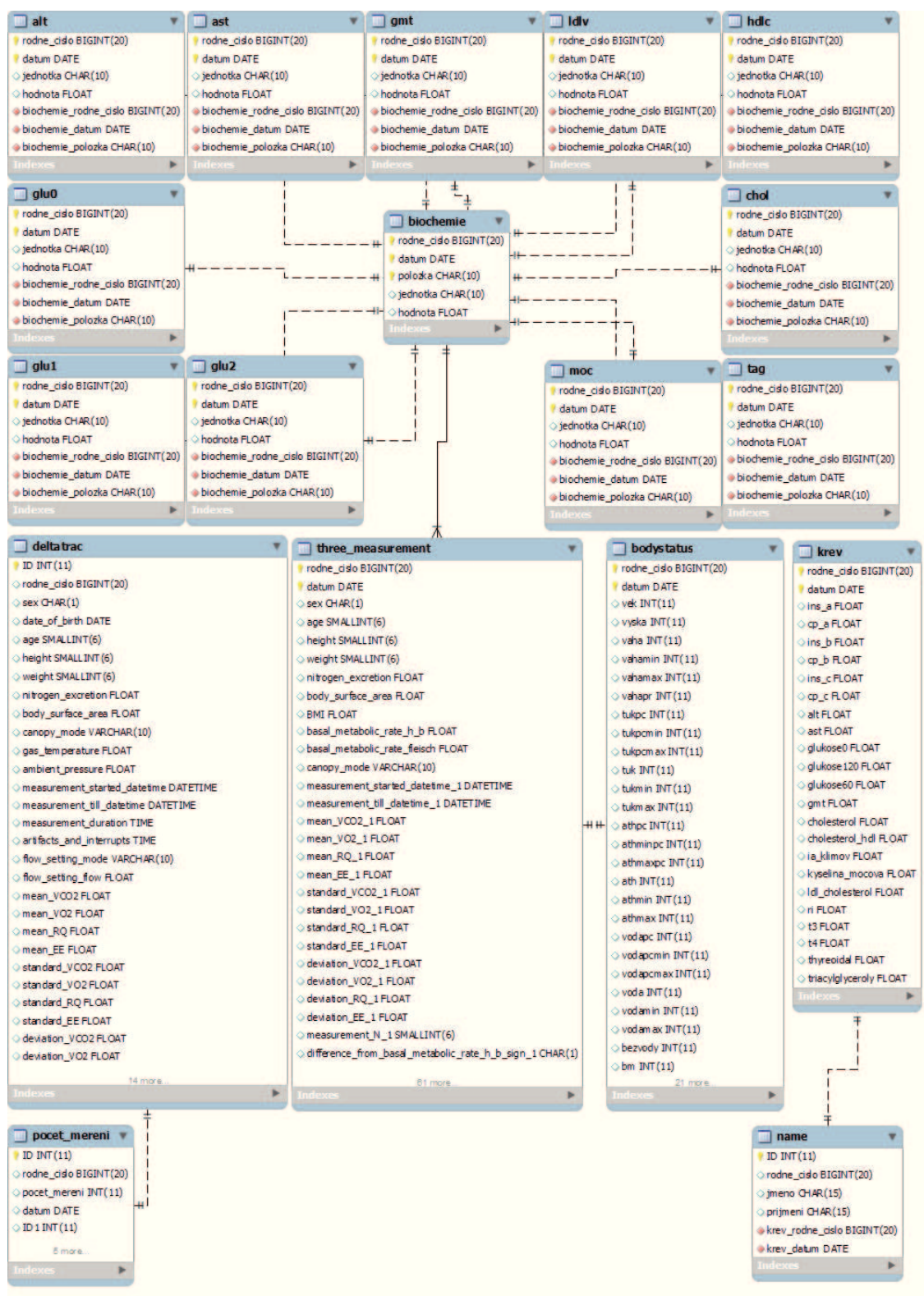
Byla použita MySQL [28] databáze z důvodů dostupnosti a jednoduchého provozování na již dostupné serverové stanici. Na obrázku 2.6 je znázorněna struktura databáze včetně všech závislostí pomocí ER diagramu [11]. Tento model se používá v softwarovém inženýrství

pro abstraktní a konceptuální znázornění dat.

Navržená databáze není v této podobě kompatibilní s elektronickou zdravotní knížkou, nicméně obsahuje všechny potřebné identifikační údaje. Pro kompatibilitu by bylo nutné navrhnout aplikaci, která provede konverzi na datový typ akceptovatelný elektronickou zdravotní knížkou.

Tabulka 2.5: Struktura databáze

Název tabulky	Počet sloupců	Počet položek	Primární klíč
deltatrac	42	11 395	ID
krev	25	931	rodne-cislo, datum
name	4	931	ID
biochemie	5	139 933	rodne-cislo, datum, polozka
bodystatus	49	271	rodne-cislo, datum
pocet_mereni	9	7 143	ID
three_measurement	84	1335	rodne-cislo, datum
alt	4	4557	rodne-cislo, datum
ast	4	4557	rodne-cislo, datum
chol	4	4418	rodne-cislo, datum
glu0	4	965	rodne-cislo, datum
glu1	4	476	rodne-cislo, datum
glu2	4	940	rodne-cislo, datum
gmt	4	4434	rodne-cislo, datum
hdlc	4	4415	rodne-cislo, datum
ldlv	4	4378	rodne-cislo, datum
moc	4	4377	rodne-cislo, datum
tag	4	4417	rodne-cislo, datum



Obrázek 2.6: ER diagram databáze

Kapitola 3

Shrnutí současného stavu

3.1 Stávající stav řešení – rešerše

Hodnocení metabolismu pacienta, potažmo hodnocení metabolických poruch a změn, je velice rozsáhlou oblastí. V tomto tématu již bylo provedeno několik primárních výzkumů. Práce v této oblasti lze rozdělit do několika skupin:

- Primární lékařské výzkumné práce
- Práce popisující chemickou bázi metabolismu
- Analytické práce na vzorku pacientů
- Klinické a souhrnné práce

Primární lékařské výzkumné práce se zabývají pochopením metabolismu z hlediska celku a fungováním lidského organismu. Vysvětlují příčiny látkové výměny, základní funkce a možné dysfunkce. Práce [8] se zabývá rozdíly metabolismu různých svalových typů a výzkum [12] řeší obecný koncept respiračního kvocientu a jeho grafickou reprezentaci. Všechny tyto práce jsou staršího data a na základě těchto analýz probíhal další výzkum v této oblasti.

Další práce popisují chemické závislosti metabolických procesů. Lze zde zmínit obecnou biochemickou literaturu [29], nebo i odborné práce [32], které popisují různé chemické příznaky poruchy metabolismu. Výzkum v těchto oblastech stále probíhá.

Analytické práce [6], [45] popisují klasifikaci jednotlivých chemických prvků nebo výsledků spektrální analýzy, nezaměřují se však na souhrnnou analýzu pacienta. Další práce [42] pak zkoumají metabolismus při různých vnějších podmínkách.

Souhrnné a klinické práce [26] shrnují poznatky z hlediska správné výživy a lékařských doporučení k dodržování režimu při různých poruchách metabolismu. Tyto publikace rovněž popisují i způsob diagnostiky metabolických poruch včetně podrobného popisu metabolických dějů.

Z hlediska poznatků v oblasti využití fuzzy klasifikací statistických souborů se vyskytují obecné publikace popisující principy klasifikace [1], [33]. Dále pak to jsou odborné publikace aplikací fuzzy klasifikací a evaluací získaných dat [20], [9].

Mezi další publikace patří zejména články aplikačního výzkumu fuzzy klasifikací na určitém typu dat a články modifikací algoritmu fuzzy klasifikace. Dále pak články evaluačních algoritmů a klasifikace medicínských dat:

- Odborná publikace [21] popisuje aplikaci fuzzy algoritmu c-průměrů pro detekci jamming stavů na základních stanicích mobilních operátorů.
- Studie [10] se zabývá fuzzy modelem neuronové sítě, která představuje Takagi-Sugeno fuzzy model.
- Publikace [2] popisuje různé implementace algoritmu fuzzy klasifikace a problémy s implementací spojené.
- Článek [23] nabízí vylepšení c-means algoritmus pro dolování dat z textových polí. Výsledný algoritmus je stabilnější a přesnější než tradiční fuzzy klasifikace.
- Odborný příspěvek [38] se zabývá fuzzy klasifikací dokumentů. Dále publikace nabízí evaluaci algoritmu dle různých indexů.
- Publikace [30] popisuje upravený algoritmus fuzzy klasifikace webových souborů včetně evaluace výsledků.
- Studie [39] se zabývá novou metodou fuzzy klasifikace s hraniční hodnotou příslušnosti. Obsahuje informace o stanovení ideálního počtu shluků a testovacích postupech.
- Článek [37] nabízí adaptivní fuzzy neuronovou síť pro klasifikaci dat. Popisuje dva algoritmy, adaptivní neuronovou síť (ANFIS) a fuzzy algoritmus (FCMDA).
- Odborný článek [44] popisuje techniky dolování dat na souboru pacientů s rakovinou prsu.
- Publikace [34] se zabývá popisem hierarchických fuzzy algoritmů pro umělá a reálná data.
- Práce [15] nabízí aplikaci fuzzy dolování dat v rozsáhlých databázích zaměstnaneckých parametrů.
- Článek [41] popisuje dolování dat pomocí fuzzy genetických metod včetně experimentálního testování.
- Studie [46] nabízí novou metodu pro fuzzy shlukovou analýzu založenou na AFS fuzzy logice.
- Práce [47] se zabývá hodnocením výsledků klasifikace pomocí separačních kritérií. Dále popisuje identifikaci dobře oddělených dat.
- Článek [14] popisuje fuzzy klasifikaci a techniky detekce pro bioinformatická data a data pacientů.
- Publikace [31] se zabývá klasifikací medicínských dat k predikci pravděpodobnosti onemocnění.
- Studie [35] nabízí aplikaci fuzzy klasifikace pro medicínské obrazy a metody dolování dat pro tyto data.

Disertační práce se zaměřuje na celkovou klasifikaci pacienta do metabolických skupin na základě analýzy dat z energometrických testů na velké skupině pacientů. Práce je unikátní právě rozsáhlostí zkoumaných parametrů a zároveň i velkým analytickým a testovacím vzorkem pacientů. Tento aplikovaný výzkum se zaměřuje na pacienty metabolické poradny, tudíž výběr je tímto zkreslen, nicméně pro zjištění metabolických poruch je tento výběr adekvátní.

3.2 Předběžná a refereční studie

Jako předběžná studie byla provedena práce [18] pod vedením autora této disertační práce, která měla za úkol částečně analyzovat data metabolického monitoru DeltaTrac II.

Cílem této práce bylo vyvinout v programovacím prostředí MATLAB software, který je schopen klasifikovat velké množství komplexních metabolických patientských dat získaných v průběhu klinických vyšetření pacientů nemocnice v Hradci Králové v letech 2007 až 2010. Program byl založen na metodách umělé inteligence a jeho jádro bylo tvořeno neuronovou sítí s architekturou self-organizing map. Aplikace dokáže zpracovat vícerozměrná vstupní data a klasifikovat je do skupin s podobnými parametry, clustery. Na základě grafických a numerických výstupů z programu pak mohou lékaři určit, které metabolické parametry navzájem souvisí a které vazby jsou naopak bezvýznamné. Veškerá patientská data jsou uložena v datábázovém systému MySQL.

K implementaci byly použity algoritmy obsažené v Neural Network Toolboxu prostředí MATLAB. Aplikace generuje grafické i numerické výsledky, na základě kterých lze interpretovat klasifikovaná data. Velkou výhodou programu je skutečnost, že se jedná o obecný algoritmus, který je možné použít nejen ke klasifikaci komplexních metabolických patientských dat, ale v podstatě ke klasifikaci kterýchkoliv vícerozměrných vstupů. Program je schopen klasifikovat data na základě až osmi parametrů. Jeho nevýhodou je fakt, že z důvodu náhodné inicializace hodnot vah mezi jednotlivými neurony modelu neuronové sítě dochází při každém použití programu k odlišným výsledkům, což může být nevyhovující při použití v deterministických systémech.

Jako testovací parametry vstupních dat byl použit stupeň obezity pacientů (BMI) a procentuální využití tuku ve všech třech měřeních přístrojem DeltaTrac II. Program klasifikoval pacienty na základě těchto parametrů do různorodých 32 shluků, což bylo otestováno statistickou analýzou.

Výsledky této práce [18] lze pro tuto disertační práci využít v několika oblastech:

1. Zpracování databáze a zkušenosti s návrhem rozsáhlé databáze
2. Výběr metody analýzy dat
3. Porovnání složitosti implementace v praxi
4. Srovnání výsledků a efektivity klasifikace

Z databázových prostředků může být využita část věnující se segmentaci a ukládání dat z měření přístrojem DeltaTrac II. Tato část musí být upravena z hlediska přizpůsobení k ostatním databázovým datům. Výběr metody analýzy dat může sloužit k následnému rozhodování o vhodnosti zpracování všech dat pomocí analýzy samoorganizující se neuronovou sítí. Zkušenosti s implementací a následnou přenositelností algoritmu do praxe jsou zásadní informací pro rozhodování o správném postupu v disertační práci. V neposlední řadě lze vzít v potaz srovnání výsledků a efektivity klasifikace při použití tohoto algoritmu nebo algoritmů jiných.

3.3 Současný stav vyhodnocení metabolického vyšetření

Vyšetření je v současné době hodnoceno dle subjektivního pohledu lékaře a jeho zkušeností, velký vliv má i anamnéza pacienta. Určení typu vyšetření závisí nejen na rozsahu diagnostického požadavku, ale i na ekonomických parametrech, jelikož vyšetření je plně hrazeno pacientem.

Neexistence standardů vyšetření vede k nejednotnosti posuzování energometrických testů a zvýšeným nákladům na hodnocení vyšetření. Práce povede ke stanovení standardů v hodnocení metabolických vyšetření a postupnému stanovení metabolických typů pacienta.

Kapitola 4

Cíl práce

4.1 Obecné cíle práce

Základním cílem práce je navrhnout algoritmus pro klasifikaci metabolismu pacienta, který umožní na základě předchozích zkušeností nebo dat stanovit jeho stav a druh poruchy metabolismu. Celkový logický postup lze shrnout do těchto bodů:

1. Rešeršní zkoumání tématu
2. Měření a zpracování potřebných dat
3. Návrh řešení algoritmu pro klasifikaci
4. Vytvoření klasifikačního vzoru
5. Testování navrženého řešení

Po základním zkoumání dosavadního řešení tématu následuje měření a zpracování potřebných dat. Po získání všech parametrů pro návrh, je navržen způsob řešení algoritmu pro klasifikaci. Tento algoritmus je nastaven dle vzoru a potom řádně otestován.

Cílem práce je tedy navržení univerzálního algoritmu pro klasifikaci, který je schopen přijmout vzor rozhodování dle předchozích zkušeností získaných z dat jiných pacientů. Tento algoritmus bude schopen do klasifikačního prostoru zařadit další pacienty, a tak provést jejich diagnostiku.

4.2 Aplikační cíle

Aplikační cíle se dají shrnout do dvou základních oblastí:

1. Primární výzkumné účely
2. Praktická aplikace v diagnostice pacienta

Algoritmus lze použít pro primární lékařské výzkumné účely, a to zejména při stanovení typologie pacienta. To je umožněno pomocí indikátoru počtu shluků v klasifikaci a souhrnných vlastností jednotlivých shluků. Výsledky tohoto primárního výzkumu lze hodnotit pomocí statistických metod.

Dále lze prakticky algoritmus uplatnit v aplikované diagnostice na lékařských pracovištích, kdy na základě výsledků primárního výzkumu je stanoven vzor, který bude sloužit jako předloha diagnostické klasifikace. Pak lze pomocí definované klasifikace zařadit nového pacienta do klasifikačního prostoru a přidělit mu souhrnné vlastnosti, které byly pozorovány u skupiny v primárním výzkumu.

Kapitola 5

Návrh a realizace řešení

5.1 Možné způsoby řešení

Práci je možno řešit několika způsoby. Klasifikaci dat lze provést:

1. Tradičními statistickými metodami
2. Fuzzy statistickými metodami
3. Neuronovými sítěmi
4. Genetickými algoritmy

Tradiční statistické metody klasifikace vytvářejí shluky objektů tak, že každý z objektů se přiřadí do právě jednoho shluku. Tento přístup je výhodný pro aplikace, kde je vyhledávána příslušnost k právě jedné skupině vlastností nebo objektů.

Fuzzy statistické metody vytvářejí strukturu klasifikace nejenom podle jednotlivých shluků, ale i podle příslušnosti prvku ke shluku. Taková klasifikace je výhodná pro zjištění míry příslušnosti ke skupině sdružených vlastností nebo jednotlivým objektům. Součástí algoritmu je zjišťování vzdálenosti od středu jednotlivých shluků.

Neuronové sítě jsou moderním přístupem ke klasifikaci, konkrétně se v těchto aplikacích využívají samoorganizující sítě. Neuronové sítě jsou schopny zjistit vazby mezi jednotlivými prvky analýzy pomocí rozsáhlého stochastického zkoumání. Jsou výhodné pro rozsáhlou analýzu dat.

Genetický algoritmus je heuristický postup motivovaný evoluční biologii. Postup je vhodný pro řešení složitých problémů, kde neexistuje exaktní algoritmus. Genetický algoritmus je poměrně implementačně složitý a náročný na výpočetní prostředky.

5.2 Navržené řešení

Pro řešení práce byla vybrána analýza pomocí fuzzy statistických metod. Je to z důvodů její snadné implementovatelnosti a přenositelnosti, dále pak umožňuje přiřazení prvků dle příslušnosti, a tak umožní zjišťovat míry jednotlivých sdružených vlastností.

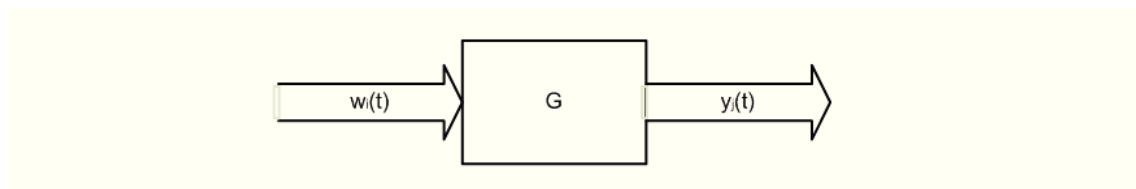
Tradiční statistické metody shlukové analýzy byly zamítnuty z důvodu nemožnosti klasifikovat pacienta do několika skupin zároveň. Rovněž neumožňují stanovit míru příslušnosti k danému shluku, což je pro koncepci práce zásadní.

Pro řešení práce by byly vhodné i metody neuronových sítí a genetických algoritmů. Ty však byly vyloučeny z důvodu obtížné implementace, výpočetní náročnosti a těžké přenositelnosti do praktických aplikací. Ze studie [18] vyplývá, že výsledky klasifikace samoorganizujícími se sítěmi nejsou vždy stejné a odvíjejí se od počátečního nastavení vstupních vah neuronové sítě. Dále z práce vyplývá, že algoritmus není jednoduše přenositelný a má velké výpočetní nároky.

Fuzzy statistické algoritmy byly vybrány, protože jejich implementace do praxe bude nejednodušší a poskytují přesný deterministický popis postupu a vlastností jednotlivých shluků.

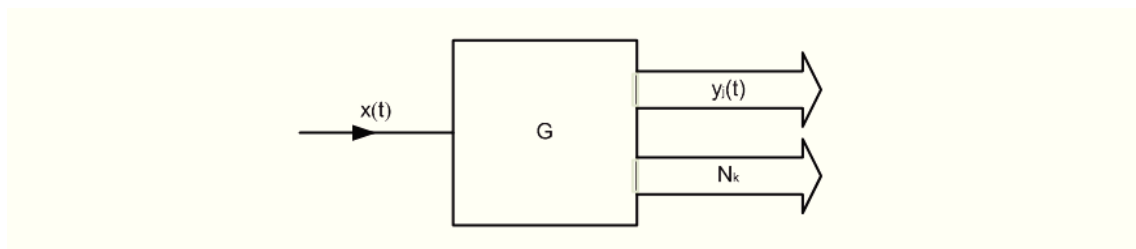
5.3 Obecný model řešení

Část procesu v lidském těle lze nazvat soustavou (dle obrázku 5.1) s přenosem G , kde $w_i(t)$ jsou vstupní signály (podněty) a $y_j(t)$ jsou výstupní signály (reakce).



Obrázek 5.1: Obecný model

Často nelze analýzu provádět pouze na základě obecných podnětů, ale je nutno aplikovat i podněty umělé, které dokážou vyvolat potřebnou reakci v organismu. Z reakce na konkrétní podnět lze pak vyvodit diagnózu.



Obrázek 5.2: Diagnostický model

Na obrázku 5.2 je uvedeno schéma modelu s diagnostickým podnětem. Soustava s přenosem G je ovlivněna vstupním signálem (podnětem) $x(t)$, výstupem jsou signály $y_j(t)$ a statický popis systému (pacienta) N_k . Právě dle výstupu $y_j(t)$ a statického popisu N_k provádíme následnou klasifikaci pacienta. Vstupním podnětem $x(t)$ se rozumí podání glukózy v definovaném čase.

Kapitola 6

Analýza dat

6.1 Výběr pacientů

Pro test bylo vybráno 254 pacientů, kteří prošli náročným výběrovým testem. Při testu byl kladen velký důraz na úplnost dat a správnou metodiku měření. Byly vybráni ti pacienti, kteří absolvovali měření typu A se správným časovým harmonogramem měření respiračního kvocientu. Pacienti rovněž museli mít kompletní biochemický rozbor krve, hodnoty insulínu, c-peptidu a změřeny fyziologické hodnoty pomocí bioimpedančního měření. Limitujícím faktorem výběrového testu byla dostupnost biochemických rozborů pouze v roce 2009, proto byl počet testovaných objektů takto omezen.

Předpokládá se, že byli vybráni pacienti:

- V širokém spektru metabolických poruch
- Zastoupení oběma pohlavími
- Se všemi hodnotami BMI
- Ve všech věkových skupinách

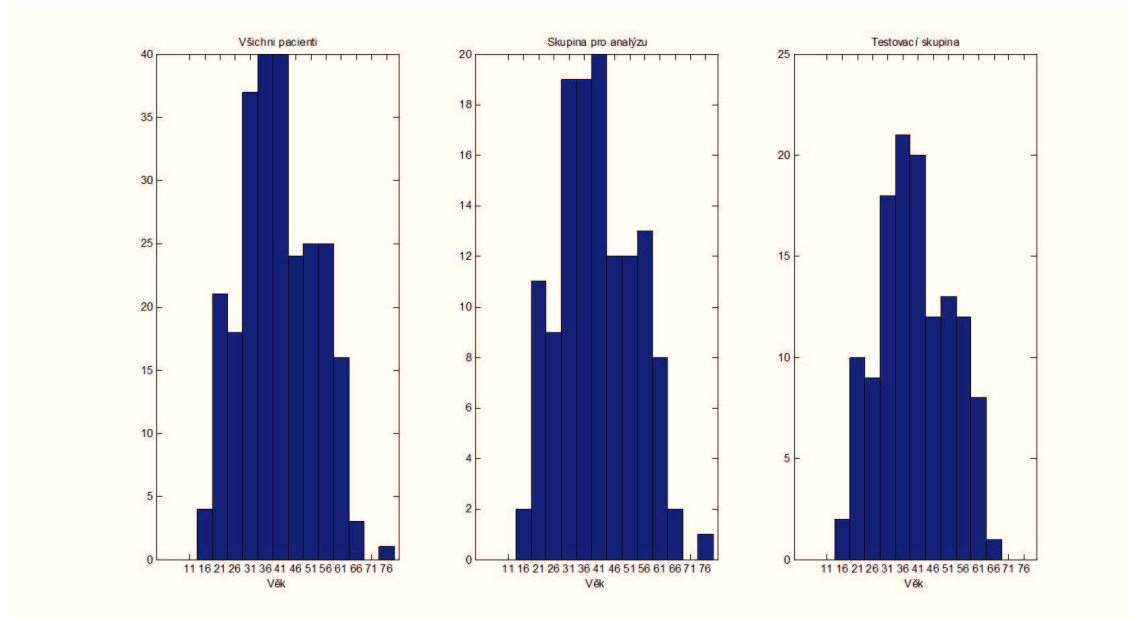
Toto se předpokládá, protože byli vybráni všichni pacienti kompletního vyšetření, kteří absolvovali návštěvu metabolické poradny během roku 2009.

6.2 Popis dat pro analýzu a testování

Vybraný soubor pacientů byl rozdělen na dvě části. Skupina pro analýzu obsahuje 128 osob a skupina pro testování algoritmu má 126 osob. Toto rozdělení umožní korektní testování algoritmu a ověří správnost navržené klasifikace. V celkovém souboru se vyskytovalo 47 mužů a 207 žen. Rozložení věku všech pacientů, skupiny pro analýzu a testovací skupiny lze vidět na obrázku 6.1.

Obrázek 6.2 popisuje rozložení věku mužů v celkovém souboru, ve skupině pro analýzu a testovací skupině. Z celkového počtu 47 mužů bylo pro analýzu použito 24 osob a pro testování osob 23. Výběr testovacích pacientů byl proveden tak, aby byly zastoupeny subjekty rovnoměrně dle věku stejně jako v analytické skupině.

Rozdělení dle pohlaví je hlavně ze statistických důvodů a preventivní korektnosti měření, aby se stejný počet mužů a žen vyskytoval jak v testování, tak v analýze.



Obrázek 6.1: Rozložení věku všech pacientů pro analýzu a testování

Rozložení věku žen ve skupině pro analýzu a testovací skupině lze vyčíst z obrázku 6.3. Z celkového počtu 207 žen bylo pro analýzu vybráno 104 osob a pro testování 103 pacientek. V testovací skupině jsou obsaženy pacientky, které reprezentují věkem skupinu pro analýzu.

Zobrazení pomocí krabicového diagramu pro všechny pacienty, pouze pro muže a pouze pro ženy je v příloze. Konkrétně se jedná o obrázky A.1, A.3 a A.2.

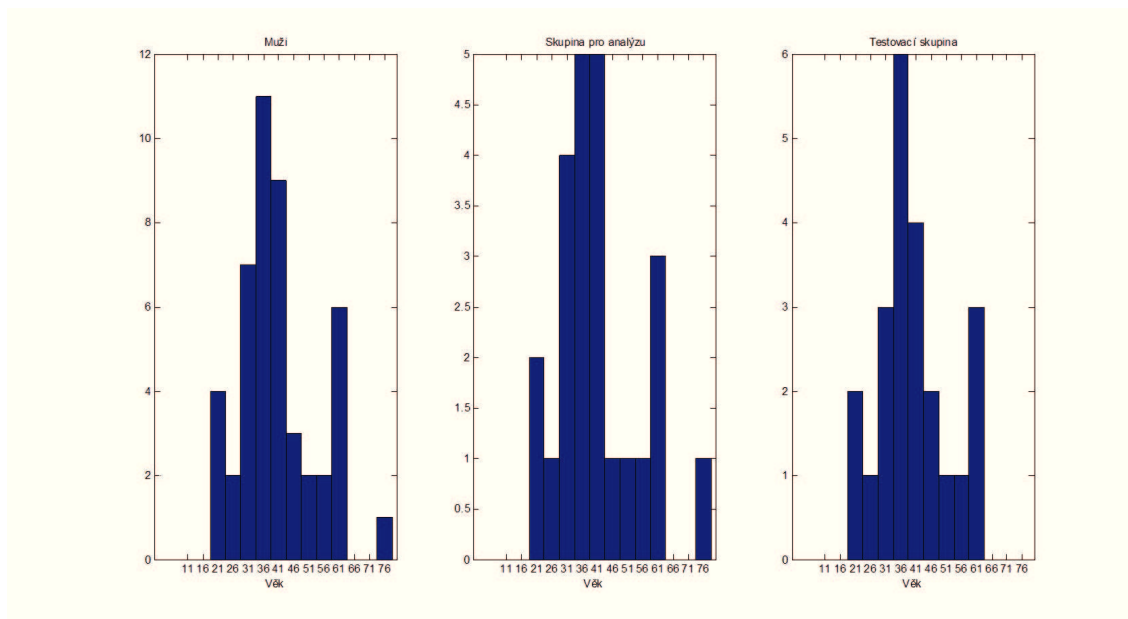
6.3 Výběr metody analýzy dat

K analýze dat před samotnou klasifikací lze použít několik postupů [17]:

1. Analýza hlavním komponent
2. Faktorová analýza
3. Korespondenční analýza
4. Diskriminantní analýza

Analýza hlavních komponent je metoda transformace sloužící k de Korelaci dat. Používá se ke snížení dimenze dat, kdy je míra ztráty informace co nejmenší. Výsledkem je soubor vektorů, hlavních komponent, které vysvětlují původní soubor dat. Tento soubor nových vektorů je získán pomocí lineární transformace původních dat. Faktorová analýza je statistická metoda, která vysvětluje rozptyl originálních dat pomocí menšího počtu latentních proměnných, které se nazývají faktory. Faktorová analýza se používá, kdy je nutné měřit něco, co není měřitelné přímým pozorováním [17].

Korespondenční analýza je metoda velice podobná analýze hlavních komponent, s tím rozdílem, že vstupní data nemohou být menší než nula. Dále používá jiné algoritmy de Korelace



Obrázek 6.2: Rozložení věku mužů pro analýzu a testování

než analýza hlavních komponent. Diskriminantní analýza se používá k predikci závislosti dat, je statistickým opakem metod ANOVA (Analysis of variance) a MANOVA (Multivariate analysis of variance) [17].

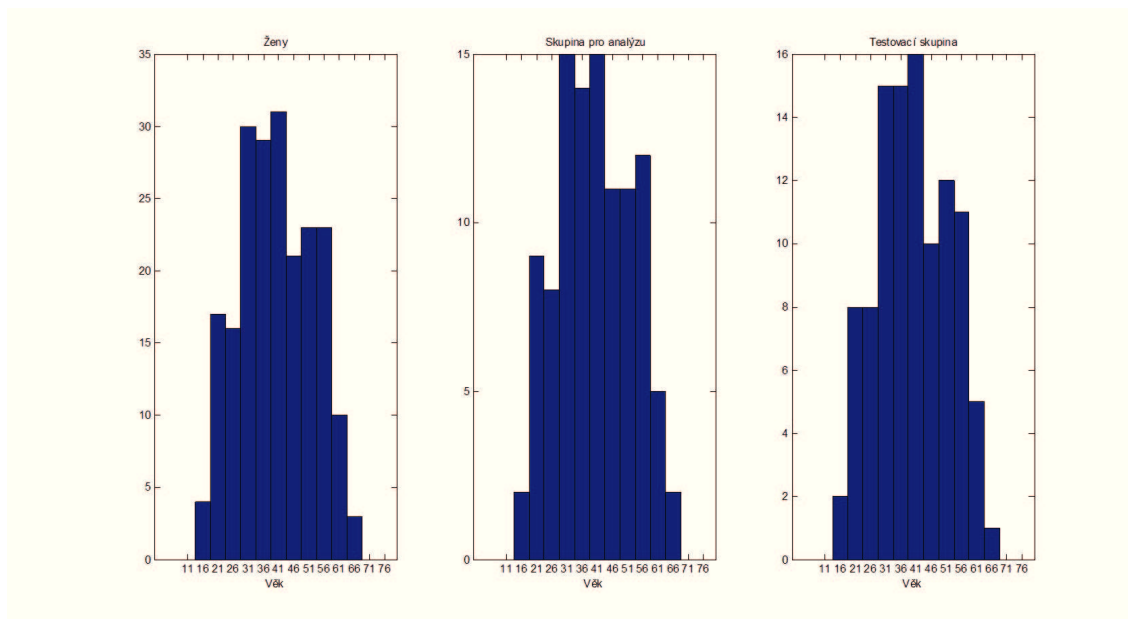
K řešení této práce je výhodné použít analýzu hlavních komponent nebo korespondenční analýzu. Byla vybrána analýza hlavních komponent z důvodu větší univerzálnosti použití na hodnotách menších než nula, dále pak z důvodu implementované funkce pro zobrazení korelačních souvislostí jednotlivých parametrů.

Analýza hlavních komponent umožňuje analyzovat jednotlivé parametry z hlediska vzájemné korelace. Transformací se získají hlavní komponenty, které reprezentují jednotlivé informační složky signálu. Součástí algoritmu je i výpočet příslušnosti jednotlivých parametrů k hlavním komponentám. Pokud různé parametry korelují s jednou hlavní komponentou, tak se předpokládá, že mají stejnou informační hodnotu a tudíž korelují i mezi sebou.

6.4 Analýza hlavních komponent

Hlavním úkolem při analýze rozsáhlých souborů dat a jejich klasifikaci je nalezení lepší reprezentace pomocí vhodné transformace. Analýza hlavních komponent (PCA – Principal Component Analysis) je klasická metoda statistické analýzy a komprese dat. Cílem této analýzy je nalézt na základě rozsáhlého počtu proměnných menší množinu nových proměnných s menší redundancí, která by poskytovala nejlepší možnou reprezentaci dat. Přitom redundance je měřena pomocí korelace mezi proměnnými. Odstranění redundance dat je stěžejní právě pro klasifikaci, kde nadbytečnost datového souboru může způsobit nečitelnost, nebo nemožnost klasifikace. Hlavní komponenty jsou umělé veličiny vytvořené jako lineární kombinace vstupních parametrů. Tyto komponenty lze vypočítat pomocí transformační matice. [33].

Analýza hlavních komponent byla provedena na skupině pro analýzu, kde bylo zařazeno



Obrázek 6.3: Rozložení věku žen pro analýzu a testování

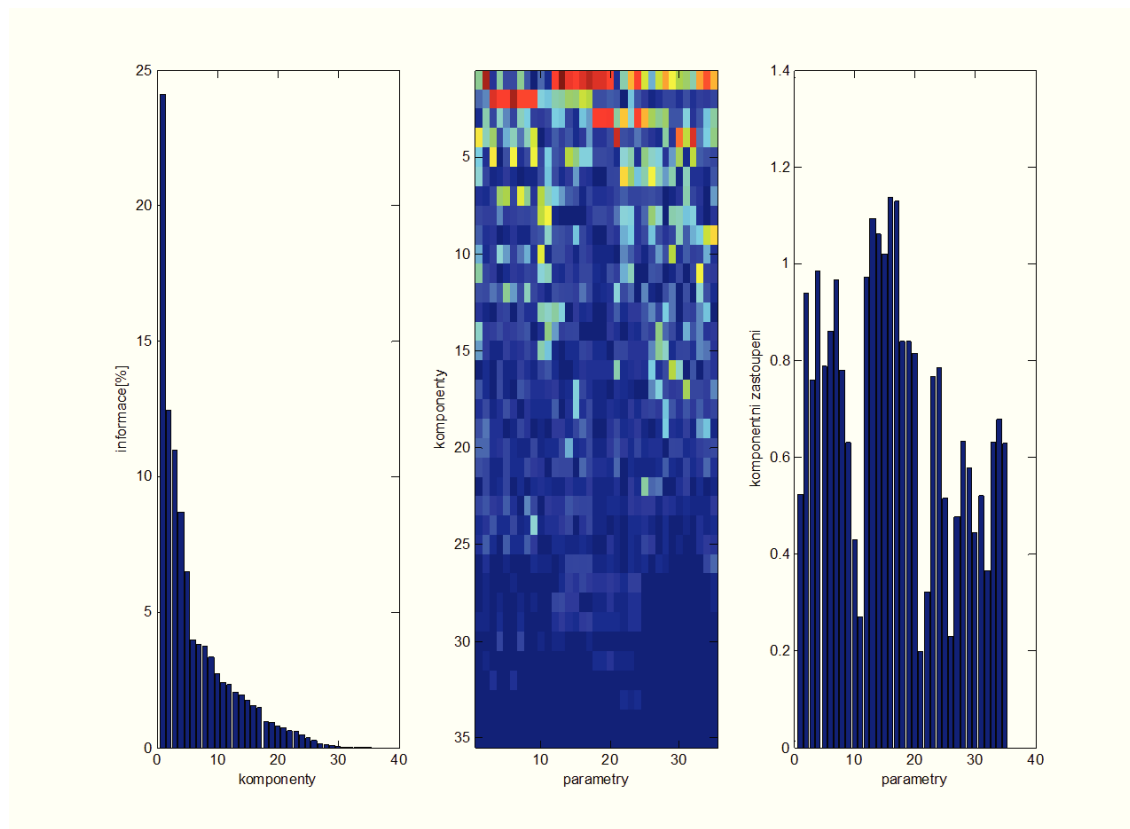
všech 35 parametrů podrobně popsanych v tabulce 6.1. Jedná se o parametry získané z měření respiračního kvocientu, konkrétně věk, body mass index (BMI), spalovaný tuk v procentech za každou hodinu testu a taktéž spalované karbohydráty. Dále se jedná o 6 parametrů měření insulinu a c-peptidu ve třech hodinových testech, hodnoty biochemického rozboru krve a parametry změřené přístrojem QuadScan 4000.

Na obrázku 6.4 je v první části zobrazena informační hodnota hlavních komponent všech dostupných parametrů. Graf zobrazuje kolik z celkové informační hodnoty parametrů je obsaženo v každé umělé vypočtené hlavní komponentě. Tento výpočet je prováděn pomocí statistické analýzy, kdy informační hodnota je posuzována podle základních parametrů rozptylu dat.

Tabulka 6.1: Parametry celkového testu

Měření	Počet parametrů	Názvy parametrů
DeltaTrac II	11	věk, BMI, RQ (3x), fat (3x), carbohydrate (3x)
Insulinové testy	6	insulin (3x), c-peptid (3x)
Biochemie	11	alt, ast, gmt, moč, tag, ldlv, hdlc, chol, glu (3x)
Bodystatus	7	tuk, ath, voda, bezvody, ict, ect, ctv
Celkem	35	

Komponenty v tomto testu jsou zastoupeny v široké škále měřených parametrů, kde prvním pěti komponentám náleží informační hodnota větší než 5 %. Z toho vychází potřeba zredukovat počet měřených parametrů, aby bylo možné použít v klasifikaci co největší informační hodnotu v prvních komponentách.



Obrázek 6.4: Informační hodnota hlavních komponent v celkovém testu

Obrázek dále v druhé části popisuje informační zastoupení jednotlivých komponent právě v jednotlivých parametrech. Barevná škála zastupuje informační korelaci v procentech, kde modrá barva vyjadřuje 0 % a červená barva 100 %. Tato grafická reprezentace je přehledná z hlediska korelace jednotlivých komponent mezi sebou, ale není vypovídající jako index důležitosti dat. Proto byla naprogramována třetí část grafu, která vyjadřuje informační součet zastoupení prvních komponent v datech, v tomto případě prvních dvou.

6.5 Návrh testovacích postupů

Existuje několik podmínek, které je nutno zohlednit při plánování klasifikačních testů:

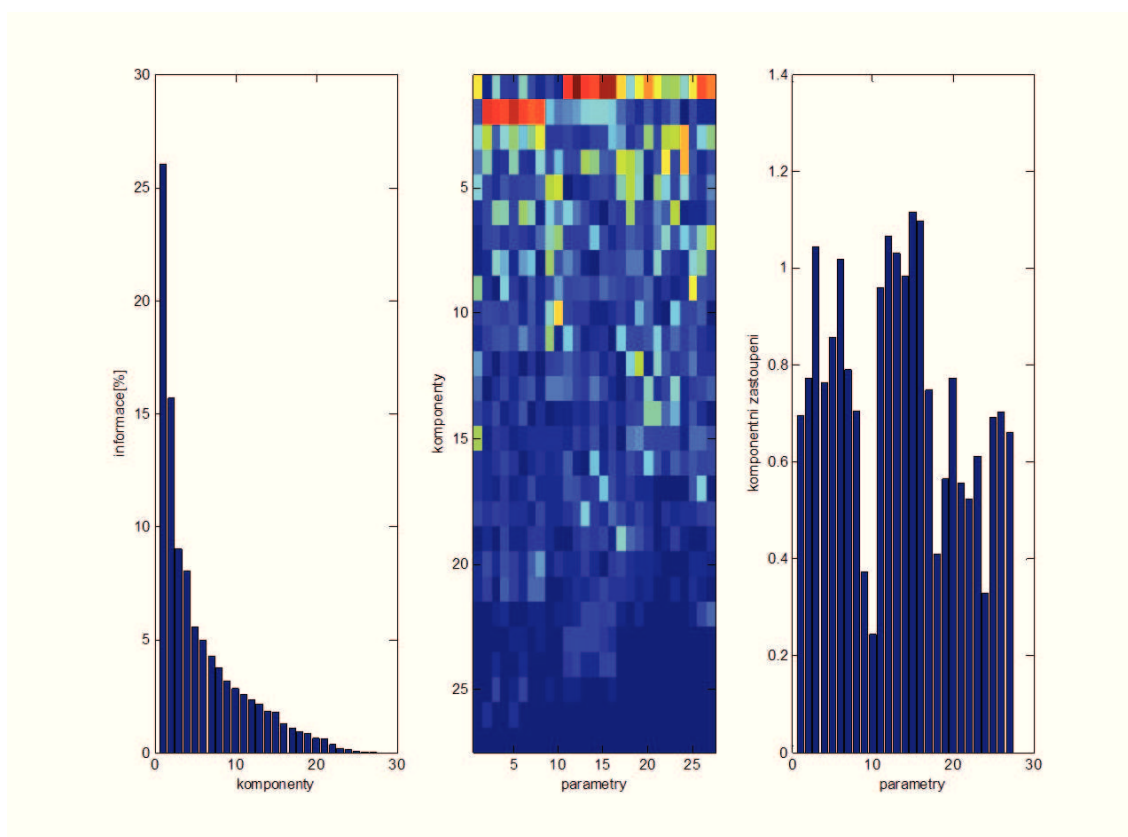
1. Počet vstupních parametrů
2. Informační hodnota hlavních komponent
3. Počet klasifikačních vektorů

Počet vstupních parametrů je zásadní pro složitost měření a následně celé klasifikace. Cílem naplnění této podmínky je minimalizace měřených parametrů na počet, který dokonale reprezentuje pacienty a nepřidává zbytečné informace pro znehodnocení klasifikace. V praxi

to znamená nalezení parametrů, které obsahují nejvíce informací v prvních hlavních komponentách.

Dalším dílčím cílem optimalizace je zvýšení informační hodnoty prvních hlavních komponent. Toto umožní podstatně zjednodušit klasifikaci a zvýšit její přesnost. Většina celkové informační hodnoty musí být soustředěna v několika prvních komponentách, toho lze dosáhnout vhodnou kombinací vstupních parametrů.

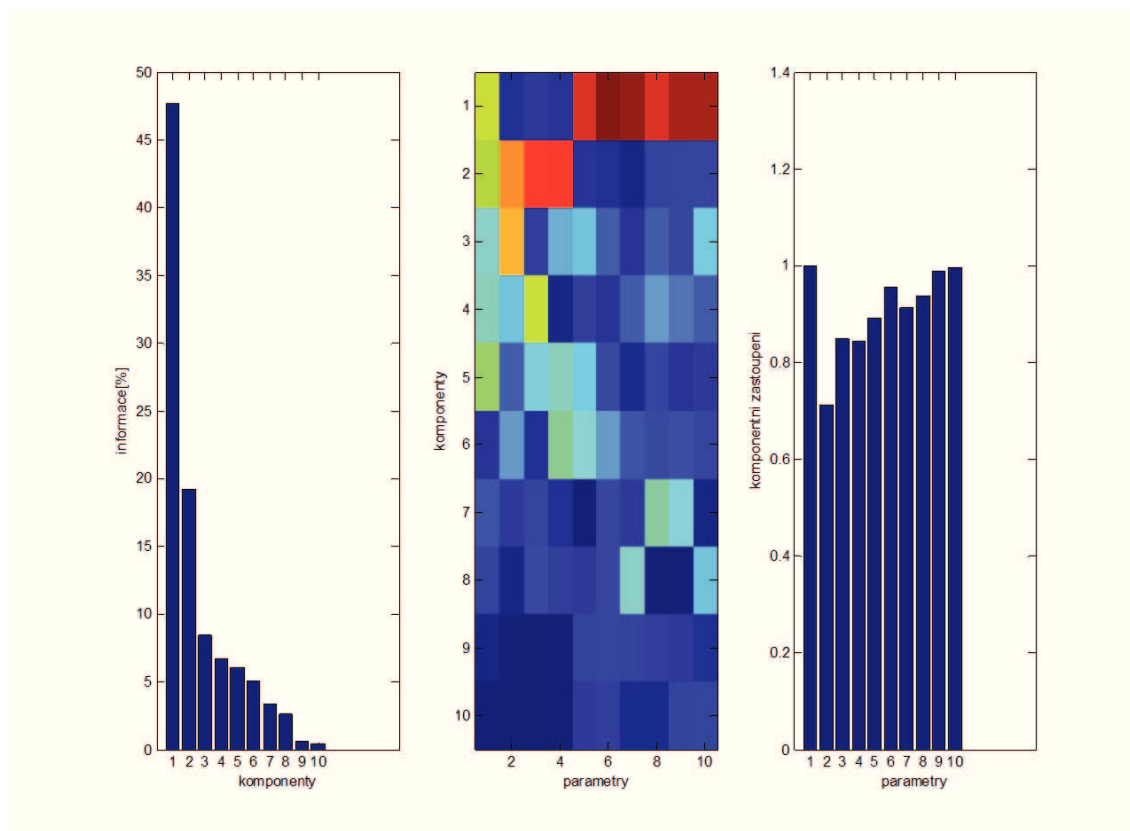
Nejvíce diskutovanou oblastí je volba počtu klasifikačních vektorů, které jsou tvořeny hlavními komponentami. Optimální případ je, když jsou voleny dvě hlavní komponenty, protože je lze názorně popsat v rovině, a tak zobrazit jednotlivé shluky. Více komponent již nemůžeme graficky zobrazit, nicméně zpřesňují výsledky klasifikace a umožňují provést klasifikaci i tam, kde je informační hodnota rozprostřena do více vektorů.



Obrázek 6.5: Informační hodnota hlavních komponent v testu A

Z těchto poznatků vyplývá další postup pro řešení práce, a to návrh tří rozdílných testů s různou volbou komponentní analýzy a klasifikace.

- Test A - DeltaTrac II a biochemie
- Test B - QuadScan 4000 a biochemie
- Test C - DeltaTrac II, QuadScan 4000 a biochemie



Obrázek 6.6: Informační hodnota hlavních komponent v redukovaném testu A

Testy A a B jsou provedeny pouze na části širokého spektra parametrů a minimalizovány tak, aby se daly zobrazit ve dvourozměrném prostoru. Test C je navržen pomocí souhrnného měření a klasifikace se provádí bez vizualizace.

6.6 Test A - DeltaTrac II a biochemie

Cílem testu A je analyzovat parametry a informační hodnotu hlavních komponent z měření respiračního kvocientu přístrojem DeltaTrac II a tuto metodu doplnit o biochemické testy a měření insulínu.

Tabulka 6.2: Parametry testu A

Měření	Počet parametrů	Názvy parametrů
DeltaTrac II	10	BMI, RQ (3x), fat (3x), carbohydrate (3x)
Insulínové testy	6	insulin (3x), c-peptid (3x)
Biochemie	11	alt, ast, gmt, moč, tag, ldlv, hdlc, chol, glu (3x)
Celkem	27	

Na obrázku 6.5 jsou v první části zobrazeny informační hodnoty jednotlivých komponent v testu A. Do testu byly zařazeny parametry uvedené v tabulce 6.2. Konkrétně se jedná o 27 parametrů, z přístroje DeltaTrac II jsou vyčítány hodnoty BMI, respiračního kvocientu, utilizace tuku a karbohydrátů. Dále jsou použita data z biochemického rozboru a insulinové testy.

Obrázek 6.5 v druhé části znázorňuje zastoupení hlavních komponent v jednotlivých testovaných parametrech. Třetí část znázorňuje součet informačních hodnot prvních dvou komponent. Pro další postup je nutné vybrat parametry pro redukovaný test tak, aby se zvýšila informační hodnota prvních dvou komponent. Existují základní tři kritéria pro výběr parametrů do redukovaného testu:

1. Redukce dle informačních kritérií
2. Redukce dle náročnosti měření
3. Redukce dle přirozené redundantnosti

Informační kritéria se posuzují dle součtu prvních hlavních komponent, toto je kritérium objektivní. Pro určení informačního kritéria byl použit koeficient 0,8, který vyjadřuje 80 % informační hodnoty v prvních dvou komponentách, tato podmínka zaručuje dostatečnou reprezentaci v klasifikační analýze. Další dvě kritéria jsou subjektivní, kdy se vyřazují parametry dle náročnosti měřeného parametru, například ekonomické nebo technické, a přirozené redundantnosti, kterou určuje lékař. Rozhodovací tabulka B.1 redukce testu A je uvedena v příloze.

Tabulka 6.3: Parametry redukovaného testu A

Měření	Počet parametrů	Názvy parametrů
DeltaTrac II	4	BMI, fat (3x)
Insulinové testy	6	insulin (3x), c-peptid (3x)
Celkem	10	

Informační hodnota hlavních komponent v redukovaném testu A je promítnuta do obrázku 6.6. Pro redukovaný test byly vybrány parametry uvedené v tabulce 6.3. Z testu byly vyřazeny hodnoty respiračního kvocientu a spalování karbohydrátů z důvodu přirozené redundantnosti, i přes těsné nesplnění informačního kritéria u parametru fat1 byl údaj zachován, protože je stěžejní pro celkové měření. Z biochemického rozboru byly odstraněny všechny parametry, protože obsahují velmi málo informací obsažených v prvních dvou komponentách, což je vidět na obrázku 6.5 v jeho třetí části, a potvrzuje to i rozhodovací tabulka B.1 uvedená v příloze této práce.

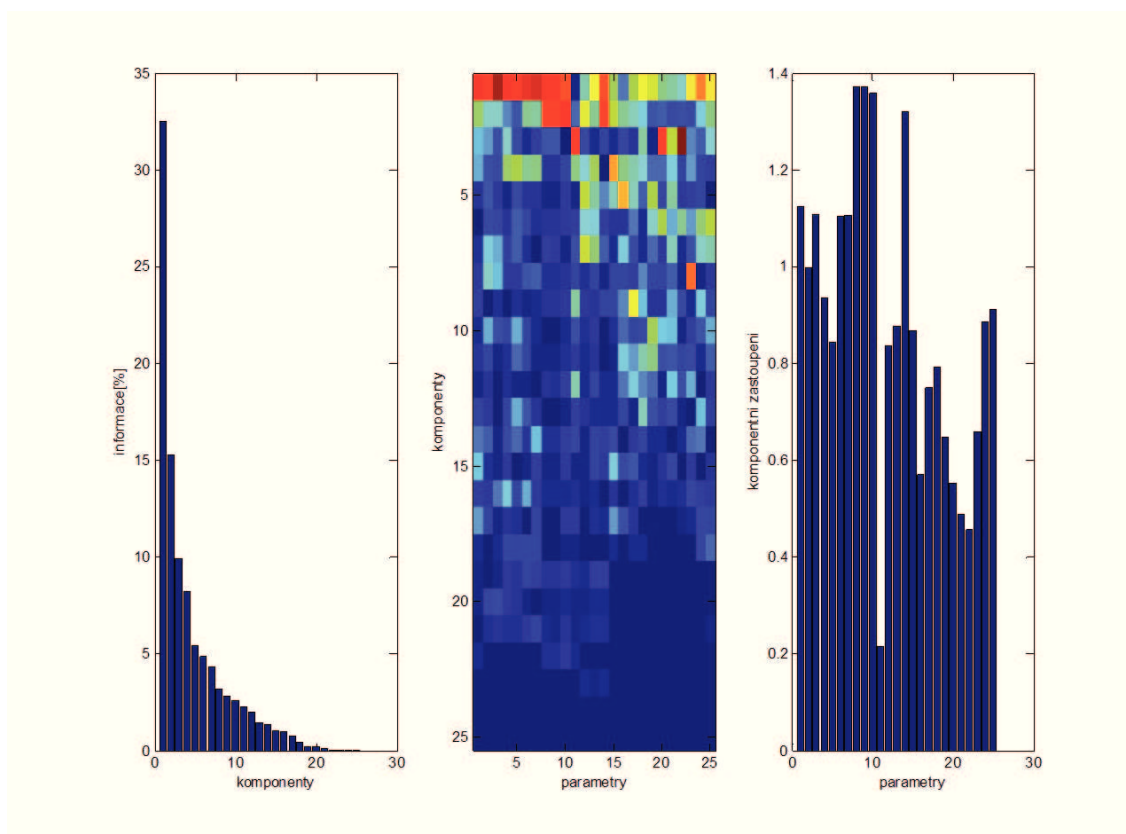
Informační hodnoty prvních dvou hlavních komponent jsou popsány tabulkou 6.4. Celková informační hodnota před redukcí je 41,69 % a po redukcí 66,87 %. Lze tedy konstatovat, že redukce přispěla ke zvýšení informační hodnoty prvních dvou hlavních komponent.

Tabulka 6.4: Informační hodnota hlavních komponent - test A

Hlavní komponenta	Informace před redukcí	Informace po redukcí
1. komponenta	26,01 %	47,67 %
2. komponenta	15,68 %	19,20 %
Celkem	41,69 %	66,87 %

6.7 Test B - QuadScan 4000 a biochemie

Test B má za úkol analyzovat parametry a informační hodnotu hlavních komponent pomocí dat měřených přístrojem QuadScan 4000. Test B bude doplněn o data biochemického rozboru krve a měření insulínu, obdobně jako v testu A.



Obrázek 6.7: Informační hodnota hlavních komponent v testu B

Graf na obrázku 6.7 v první části zobrazuje informační hodnoty jednotlivých komponent, které jsou použity v testu B. Jednotlivé parametry použité pro analýzu v testu B jsou uvedeny v tabulce 6.5. Jedná se o 25 parametrů, a to konkrétně 8 hodnot měřených systémem QuadScan 4000, 6 hodnot získaných insulínovými testy a 11 hodnot z biochemického rozboru.

Obrázek 6.7 části druhé popisuje zastoupení hlavních komponent v jednotlivých paramet-

Tabulka 6.5: Parametry testu B

Měření	Počet parametrů	Názvy parametrů
QuadScan 4000	8	BMI, tuk, ath, voda, bezvody, ict, ect, ctv,
Insulinové testy	6	insulin (3x), c-peptid (3x)
Biochemie	11	alt, ast, gmt, moč, tag, ldlv, hdlc, chol, glu (3x)
Celkem	25	

rech. Poslední část vyjadřuje součet informačních hodnot prvních dvou hlavních komponent.

Tabulka 6.6: Parametry redukovaného testu B

Měření	Počet parametrů	Názvy parametrů
QuadScan 4000	7	BMI, tuk, ath, voda, ict, ect, cvt
Insulinové testy	6	insulin (3x), c-peptid (3x)
Biochemie	3	alt, glu1, glu2
Celkem	16	

Výběr parametrů pro redukci probíhal dle stejných kritérií jako u testu A, a to konkrétně:

1. Redukce dle informačních kritérií
2. Redukce dle náročnosti měření
3. Redukce dle přirozené redundantnosti

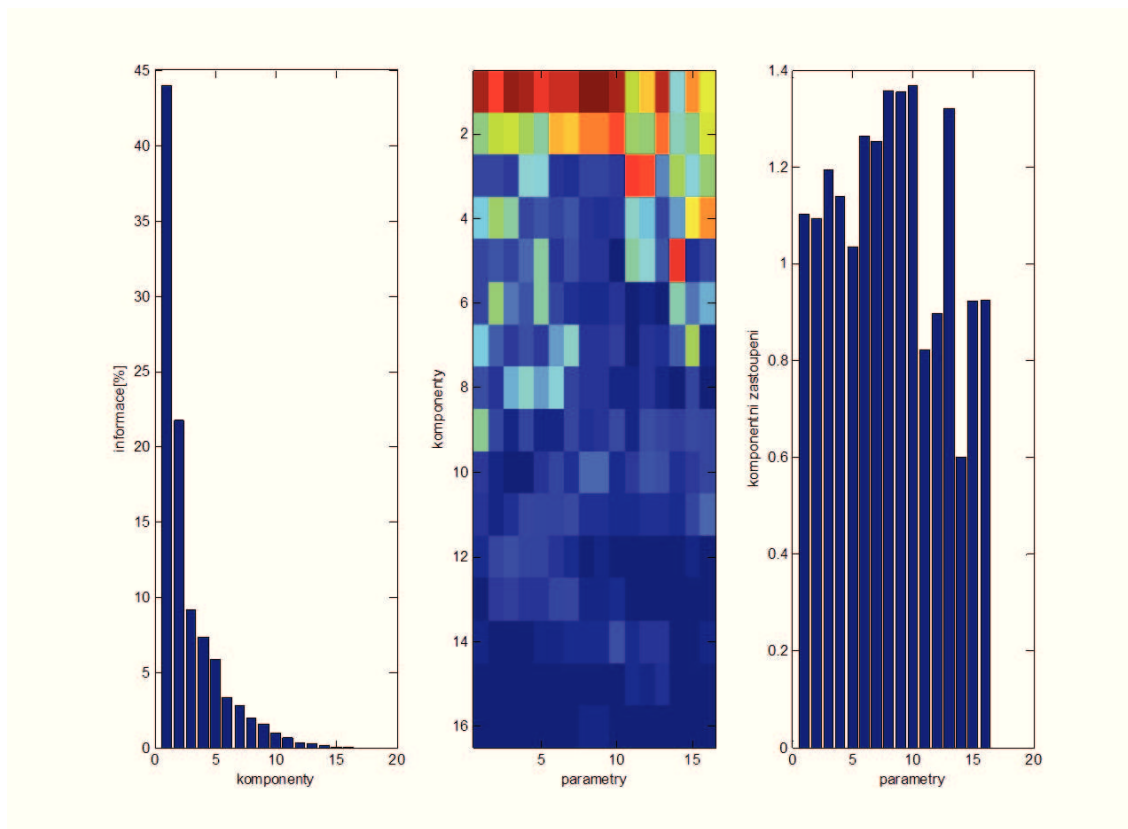
Rozhodovací tabulku B.2 lze nalézt v příloze této práce. Obdobně jako v testu A bylo použito informační kritérium 0,8, které vyjadřuje 80 % informace obsažených v prvních dvou hlavních komponentách.

Obrázek 6.8 v první části popisuje informační hodnotu hlavních komponent v redukovaném testu B. Redukovaný test byl proveden s parametry uvedenými v tabulce 6.6. V redukovaném testu B je obsaženo 16 parametrů měření. Byly částečně odstraněny parametry změřené přístrojem QuadScan 4000 a převážná většina biochemického rozboru krve. Všechny zmíněné parametry byly odstraněny z důvodu informační nedostatečnosti.

Tabulka 6.7: Informační hodnota hlavních komponent - test B

Hlavní komponenta	Informace před redukcí	Informace po redukcí
1. komponenta	32,51 %	44,01 %
2. komponenta	15,31 %	21,71 %
Celkem	47,82 %	65,72 %

Celková informační hodnota prvních dvou hlavních komponent je popsána tabulkou 6.7. Informační hodnota neredukovaného testu je 47,82 % a po redukcí je tato informační hodnota



Obrázek 6.8: Informační hodnota hlavních komponent v redukovaném testu B

zvýšena na 65,72 %. Redukce tedy opět přispěla ke zvýšení informační hodnoty prvních dvou komponent.

6.8 Test C - DeltaTrac II, QuadScan 4000 a biochemie

Test C je navržen jako souhrnný test obsahující parametry ze všech dostupných měření. Součástí analýzy jsou data měřená systémem DeltaTrac II, dále přístrojem QuadScan 4000 a kompletním biochemickým rozбором krve doplněným o informace insulínu a c-peptidu.

Na obrázku 6.9 jsou nejprve zobrazeny informační hodnoty jednotlivých komponent v testu C. Parametry použité v testu jsou souhrnně uvedeny v tabulce 6.8. Celkový počet parametrů je 34 a jsou to údaje měření respiračního kvocientu, data z měření impedance lidské tkáně a kompletní biochemická data.

Obrázek 6.9 ve střední části popisuje zastoupení hlavních komponent v jednotlivých parametrech. Opět modrá barva reprezentuje nulovou korelaci a červená barva korelaci maximální. V poslední části je uveden součet prvních 4 komponent analýzy. Tento počet komponent byl zvolen jako maximální možný a efektivní pro klasifikaci. Klasifikační algoritmus je právě schopný zpracovat až 4 vstupní vektory, aby nedošlo ke zkreslení klasifikace a zároveň informační hodnota nižších komponent je natolik malá, že není potřeba je do klasifikace zahrnovat. Komponenty s nižší informační hodnotou často obsahují šum a další nežádoucí informace.

Tabulka 6.8: Parametry testu C

Měření	Počet parametrů	Názvy parametrů
DeltaTrac II	10	BMI, RQ (3x), fat (3x), carbohydrate (3x)
QuadScan 4000	7	tuk, ath, voda, bezvody, ict, ect, ctv,
Insulinové testy	6	insulin (3x), c-peptid (3x)
Biochemie	11	alt, ast, gmt, moč, tag, ldlv, hdlc, chol, glu (3x)
Celkem	34	

Tabulka 6.9: Parametry redukovaného testu C

Měření	Počet parametrů	Názvy parametrů
DeltaTrac II	4	BMI, fat (3x)
QuadScan 4000	7	tuk, ath, voda, bezvody, ict, ect, ctv,
Insulinové testy	6	insulin (3x), c-peptid (3x)
Biochemie	11	alt, ast, gmt, moč, tag, ldlv, hdlc, chol, glu (3x)
Celkem	28	

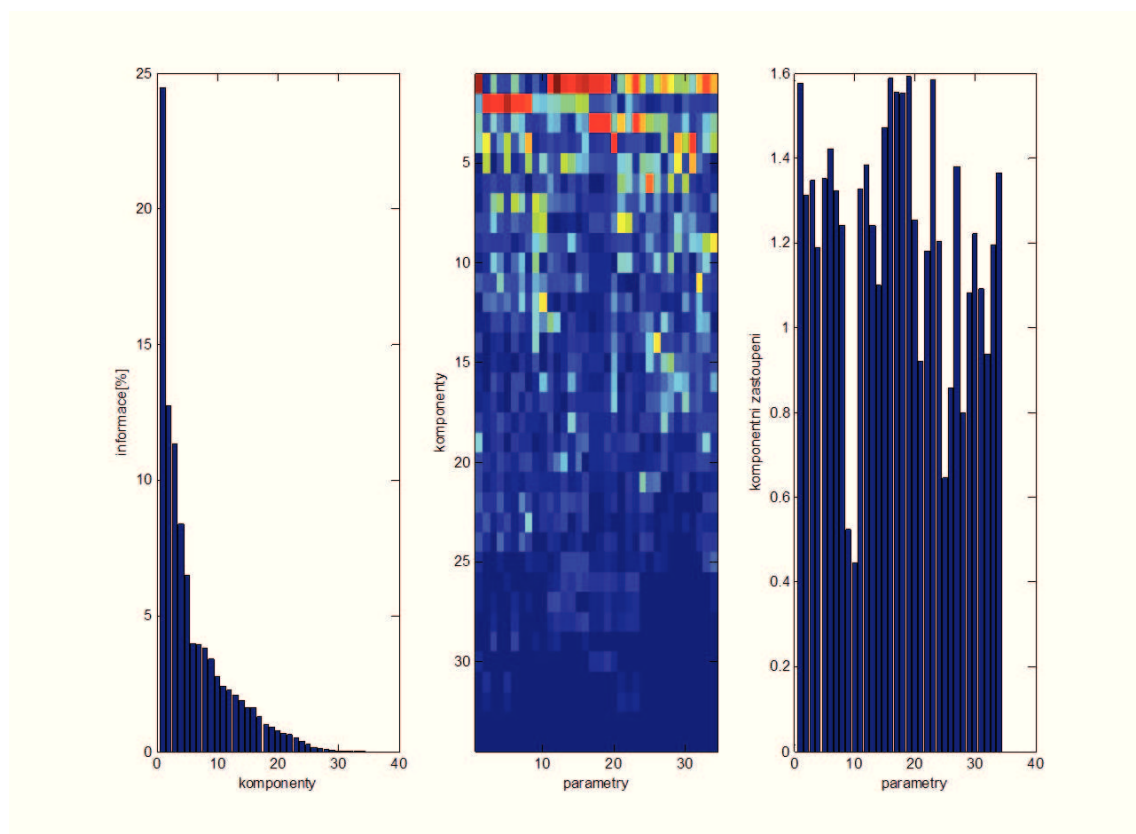
Aby bylo možno zvýšit informační hodnotu prvních komponent, musí být provedena redukce. Informační kritérium bylo v tomto případě zvýšeno na 1, což odpovídá 100 % zastoupení informace v prvních 4 komponentách. Rozhodovací tabulka B.3 je uvedena v příloze. Redukce byla provedena dle informačního kritéria a byly odstraněny redundantní informace u měření respiračního kvocientu přístrojem DeltaTrac II.

Informační zastoupení v redukovaném testu je zobrazeno na obrázku 6.10 a tabulka 6.9 popisuje veškeré parametry použité v tomto testu. Při redukci byly odstraněny pouze zdrojové parametry respiračního kvocientu pro výpočet spalování tuku a hodnoty spalování karbohydrátů, které jsou dopočteny přes spalování tuku.

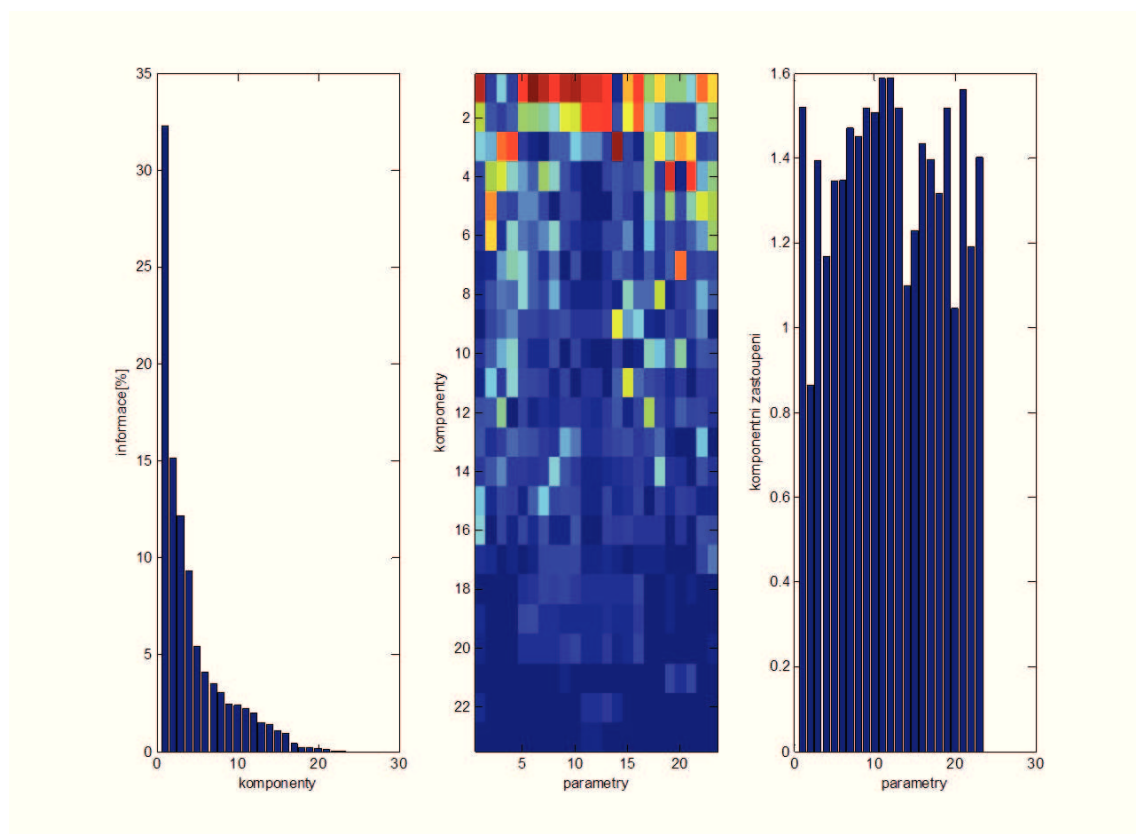
Tabulka 6.10: Informační hodnota hlavních komponent - test C

Hlavní komponenta	Informace před redukcí	Informace po redukcí
1. komponenta	24,49 %	32,29 %
2. komponenta	12,74 %	15,15 %
3. komponenta	11,33 %	12,16 %
4. komponenta	8,38 %	9,33 %
Celkem	56,94 %	68,93 %

Tabulka 6.10 popisuje hodnoty prvních 4 hlavních komponent a jejich celkový součet před redukcí a po ní. Celková informační hodnota před redukcí je 56,94 % a po redukcí se zvýšila na 68,93 %. Lze z toho vyvodit, že odstranění redundantních údajů vedlo k nárůstu informační hodnoty ve zvolených hlavních komponentách.



Obrázek 6.9: Informační hodnota hlavních komponent v testu C



Obrázek 6.10: Informační hodnota hlavních komponent v redukovaném testu C

Kapitola 7

Klasifikace dat

Základním cílem shlukové analýzy je zařadit objekty do skupin (shluků), a to především tak, aby si dva objekty stejného shluku byly více podobné, než dva objekty z různých shluků. Přitom objekty mohou být různého charakteru [33].

7.1 Tradiční metody klasifikace

Tradiční metody se dají rozdělit do dvou základních skupin, kterými jsou metody rozkladu a hierarchické metody. Hierarchické postupy jsou jasně vymezené a dochází pomocí nich k vytvoření hierarchie shluků. V metodách rozkladu dochází k zařazení objektu do předem stanoveného počtu shluků [33].

Metoda k-průměrů se používá pro tradiční klasifikaci, výpočet se provádí dle vzorce 7.1 [1].

$$J(\mathbf{X}; \mathbf{V}) = \sum_{i=1}^c \sum_{k \in i} \|\mathbf{x}_k^{(i)} - \mathbf{v}_i\|^2 \quad (7.1)$$

Kde $J(\mathbf{X}; \mathbf{V})$ je funkce střední kvadratické odchylky, určená k minimalizaci, s parametry vstupních dat \mathbf{X} a shluků \mathbf{V} . Parametr k představuje počet shluků, $\mathbf{x}_k^{(i)}$ je k -tá skupina vstupních dat patřící do i -tého shluku a \mathbf{v}_i je vektor průměrů i -tého shluku [16].

Dalším algoritmem pro tradiční shlukovou analýzu je metoda k-medoidů (PAM – Partitioning Around Medoids). Výpočet se provádí dle vzorce 7.2 [33].

$$f = \sum_{i=1}^n D(\mathbf{x}_i, \mathbf{m}_{g,i}) \quad (7.2)$$

Datové soubory jsou rozděleny do g shluků, kde jsou určeny medoidy, což jsou jednotlivé objekty ze shluku. Výpočet $D(\mathbf{x}_i, \mathbf{m}_{g,i})$ reprezentuje vzdálenost shlukovaných dat \mathbf{x}_i a medoidu $\mathbf{m}_{g,i}$ [33].

7.2 Fuzzy metody klasifikace

Fuzzy shluková analýza obsahuje míry příslušnosti jednotlivých dat k danému shluku, které tradiční metody shlukové analýzy neobsahují [33].

Tato analýza byla vybrána právě pro tuto aplikaci, protože obsahuje zmíněné míry příslušnosti k jednotlivým shlukům. Je tedy potom možné určovat typy pacientů dle příslušnosti k jednotlivým shlukům vlastností a určovat tak míru příslušnosti jednotlivé vlastnosti.

Vzorcem 7.3 je popsán výpočet c-průměrů fuzzy shlukové analýzy [1].

$$J(\mathbf{X}; U, V) = \sum_{i=1}^c \sum_{k=1}^N (\mu_{i,k})^m \|\mathbf{x}_k - \mathbf{v}_i\|_{\mathbf{A}}^2 \quad (7.3)$$

Celý výpočet je založen na minimalizaci funkce $J(\mathbf{X}; U, V)$, kde \mathbf{X} je maticí vstupních dat, U je fuzzy příslušnost jednotlivých dat a V je vektor jednotlivých shluků. Parametr k představuje počet shluků, \mathbf{x}_k je k-tá skupina vstupních dat patřící do i-tého shluku, \mathbf{v}_i je vektor průměrů i-tého shluku, $(\mu_{i,k})^m$ je stupeň příslušnosti dat ke shluku a \mathbf{A} je norma k výpočtu vzdálenosti [1].

Fuzzy klasifikace má využití v této práci, protože umožňuje klasifikované prvky zařadit do více shluků. Součástí algoritmu je tedy i míra příslušnosti k těmto shlukům. Míra příslušnosti umožní popsat pacienta, jako příslušníka několika metabolických skupin s jasnou měrou příslušnosti.

7.3 Normalizace dat

Normalizace shlukovaných dat je nutná k vyvážení poměrů dat s různými statistickými vlastnostmi. Lze tak docílit toho, že každý rozměr vícerozměrných klasifikovaných dat bude mít v klasifikaci stejnou váhu.

Práce využívá dvou základních algoritmů normalizace, je to normalizace dle variance statistického souboru a normalizace dle rozsahu statistického souboru [7].

$$X = \frac{X_{old} - X_{min}}{X_{max} - X_{min}} \quad (7.4)$$

Vzorec 7.4 popisuje první typ normalizace dat, a to základní normalizaci na rozsah statistického souboru. Ve výpočtu je aktuální hodnota ponížena o minimální hodnotu souboru, toto celé je děleno rozdílem maximální a minimální hodnoty. Základní vlastnost této normalizace je taková, že všechny hodnoty statistického souboru jsou v rozsahu 0 až 1. Tato normalizace je výhodná pro objektivizaci zobrazení v prostoru a normalizování dat před samotnou klasifikací.

Pokud je potřeba zařadit do stejného normalizovaného souboru další hodnoty, je nutné znát maximální a minimální hodnotu původního statistického souboru. Zařazení do stejného normalizovaného prostoru je nutné právě pro funkci navrženého algoritmu. Pokud však nová hodnota není v mezích původní maximální a minimální hodnoty, pak může dojít k zařazení mimo normalizovaný prostor, je tedy přiřazena hodnota mimo rozsah 0 až 1.

$$X = \frac{X_{old} - \bar{X}}{\sigma_x} \quad (7.5)$$

Druhým algoritmem popsaným vzorcem 7.5 je normalizace dle variance. Normalizovaná hodnota je vypočtena jako rozdíl původní a střední hodnoty, jenž je vydělen variancí statistického souboru. Základním rysem této normalizace je, že zohledňuje především charakter rozložení statistického souboru a neovlivní ji odlehlá pozorování, popřípadě chyby měření

mimo standardní rozsah hodnot. Tento algoritmus je výhodný pro použití před analýzou hlavních komponent, kdy je velká váha právě na objektivním posouzení vstupů.

Aby bylo možné provést normalizaci i pro dodatečné hodnoty, je nutné znát střední hodnotu a varianci původního statistického souboru. Dodatečné zařazení je opět nutné pro celý algoritmus navržený v této práci, protože takto je možné provést transformaci dle koeficientů analýzy hlavních komponent i pro nové pacienty se stejnou váhou vstupních parametrů.

7.4 Návrh počtu skupin

Návrh počtu klasifikačních skupin neboli shluků je klíčový pro celý klasifikační algoritmus. Je jednodušší určit počet shluků u souborů, kde je předběžně zřejmé, jaké typy dat jsou na vstupu a jaké jsou hlavní kategorie dat. U multiparametrické analýzy toto zřejmé není a musí být využita řada koeficientů predikující vhodnost určitého počtu shluků. Žádný z následujících vzorců není vypovídající sám o sobě, správného výsledků lze docílit pouze vzájemnou komparací jednotlivých indexů [7].

$$PC(c) = \frac{1}{N} \sum_{i=1}^c \sum_{j=1}^N (\mu_{ij})^2 \quad (7.6)$$

Prvním koeficientem je koeficient rozdělení (Partition Coefficient – PC), který je vypočten vzorcem 7.6. Parametr μ_{ij} definuje příslušnost dat s pořadovým číslem j k shluku i . Když koeficient PC stoupá, tak je daná shluková aproximace vhodnější [7].

$$CE(c) = -\frac{1}{N} \sum_{i=1}^c \sum_{j=1}^N \mu_{ij} \cdot \log(\mu_{ij}) \quad (7.7)$$

Dalším koeficientem k popisu vhodnosti shlukové aproximace je klasifikační entropie (Classification Entropy – CE) vypočtená vzorcem 7.7. Tento koeficient je velice podobný parametru PC, nicméně v optimálním počtu shluků dosahuje minima [7].

$$SC(c) = \sum_{i=1}^c \frac{\sum_{j=1}^N (\mu_{ij})^2 \|x_j - v_i\|^2}{N_i \sum_{k=1}^c \|v_k - v_i\|^2} \quad (7.8)$$

Index rozdělení (Partition Index – SC) je vypočten dle vzorce 7.7. Je to poměr kompaktnosti a separace shluků, dá se vhodně využít, když se porovnávají různá rozdělení se stejným počtem shluků. Čím je nižší hodnota tohoto indexu, tak je rozdělení lepší [7].

$$S(c) = \frac{\sum_{i=1}^c \sum_{j=1}^N (\mu_{ij})^2 \|x_j - v_i\|^2}{N \min_{i,k} \|v_k - v_i\|^2} \quad (7.9)$$

Ve vzorci 7.9 je uveden výpočet separačního indexu (Separation Index – S), který je opakem indexu rozdělení (SC). Tento index používá minimální vzdálenosti, aby určil správnost shlukové aproximace [7].

$$XB(c) = \frac{\sum_{i=1}^c \sum_{j=1}^N (\mu_{ij})^2 \|x_j - v_i\|^2}{N \min_{i,j} \|x_j - v_i\|^2} \quad (7.10)$$

Obdobným ukazatelem je i Xie a Beniho index (Xie and Beni's Index – XB) uvedený ve vzorci 7.10. Tento koeficient kvantifikuje poměr celkové variace shluků a separace shluků [7].

Poslední dva ukazatelé jsou Dunnův index (Dunn's Index – DI) a alternativní Dunnův index (Alternative Dunn Index – ADI). Výpočet těchto indexů je poměrně složitý a je podrobně popsán v literatuře [7]. Tyto ukazatele slouží k identifikaci dobré separace jednotlivých shluků.

Analýza pomocí všech těchto indexů byla provedena na testech A, B i C. Výsledky jsou uvedeny v příloze této práce a testy byly provedeny pro různé počty shluků, v tomto případě pro 2 až 14 shluků. Obecně se očekává, že aproximace větším počtem shluků bude vždy lepší, když se vyřadí lokální výkyvy. Cílem je tedy nalezení minimálního počtu shluků, které splňují požadavky několika parametrů a jsou ideálně charakterizovány lokálním nebo globálním extrémem. Je pravděpodobné, že ke stanovenému počtu shluků bude existovat i lepší vyšší počet shluků s přesnější aproximací, který však nereprezentuje cíl práce, a to je sdružení pacientů do malého počtu skupin s charakteristickými vlastnostmi.

Výsledky testu A jsou zobrazeny na obrázcích C.1, C.2 a C.3. Ukazatel PC nabývá lokálního minima v hodnotě 3 a globálního minima v hodnotě 6, hodnotu 3 potvrzuje i ukazatel CE a částečně potvrzuje i hodnotu 6. Index SC vykazuje strmý pád do hodnoty 3 a pak je pokles do hodnoty 6 pozvolný. U trendů S a XB se vyskytuje lokální maximum u hodnoty 4, což ukazuje nevhodnou aproximaci při této hodnotě, hodnoty 3 a 6 zde vykazují podobné charakteristiky. Test DI a ADI je v tomto neprůkazný, oba testy se výrazně liší a to pravděpodobně z důvodu fuzzy překrývání jednotlivých shluků. Ze souhrnného hodnocení vycházejí počty shluků 3 a 6 podobně, proto je pro test A volen počet shluků 3, neboť umožní souhrnné hodnocení větší skupiny pacientů a má menší diversifikaci souhrnných vlastností.

Test všech koeficientů pro variantu B je zobrazen na obrázcích C.4, C.5 a C.6. Koeficienty PC a CE se projevují stejně, a to tak, že mají lokální extrémy v počtu 6, zbytek funkce je monotónní. Index SC klesá prudce k hodnotě 3 a pak klesá pozvolna k hodnotě 9. Podobně se chová i separační index S, který v oblasti 3 až 5 je konstantní a pak dosahuje globálního extrému v hodnotě 7. Index XB nekoresponduje se žádným předchozím indexem. Ukazatel DI preferuje hodnotu 7, stejně jako index S, ale ukazatel ADI je neprůkazný. V tomto případě je rozhodování velice složité, protože žádný index jasně neurčil správný počet shluků.

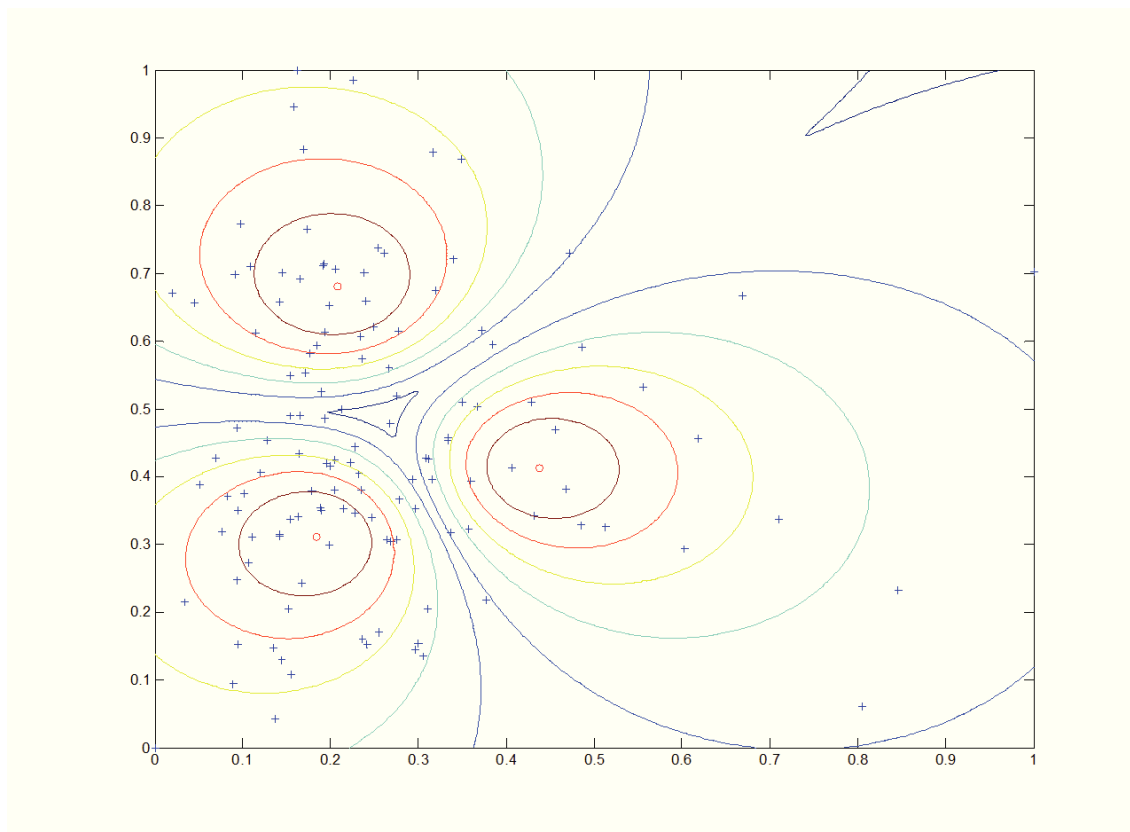
Ukazatele pro test C jsou zobrazeny na obrázcích C.7, C.8 a C.9. Koeficienty PC a CE jsou do hodnoty 11 monotónní, takže jsou pro detekci ideálního počtu shluků neprůkazné. Indexy SC, S a XB strmě klesají k hodnotě 3 a pak klesají zvolna, index XB má lokální minimum na hodnotě 11. Ukazatelé DI a ADI se chovají rozdílně, index DI preferuje hodnotu 3. Z celkového srovnání indexů je zvolena hodnota 3 pro nejlepší aproximaci.

Z analýzy vyšlo, že pro testy A a C je nejlepší klasifikace pomocí 3 shluků, test B byl neprůkazný. Při předpokladu, že testy jsou tvořeny stejným statistickým vzorkem pacientů s podobnými vlastnostmi testovaných parametrů, byla i testu B přiřazena hodnota 3 shluků. Všechny 3 testy budou tedy klasifikovány 3 shluky a následně otestovány, zda tato aproximace je vyhovující.

7.5 Klasifikační algoritmus

Klasifikace je provedena fuzzy algoritmem, který v každém testu zařadí pacienta do shluků s patřičným koeficientem příslušnosti. U všech testů byla zvolena klasifikace do tří shluků.

Na obrázku 7.1 je zobrazena klasifikace testu A provedeného pomocí dvou hlavních komponent, v každém shluku je umístěn centroid, který reprezentuje geometrický střed shluku. Data před testem jsou normalizována algoritmem range, který data normuje v celém rozsahu

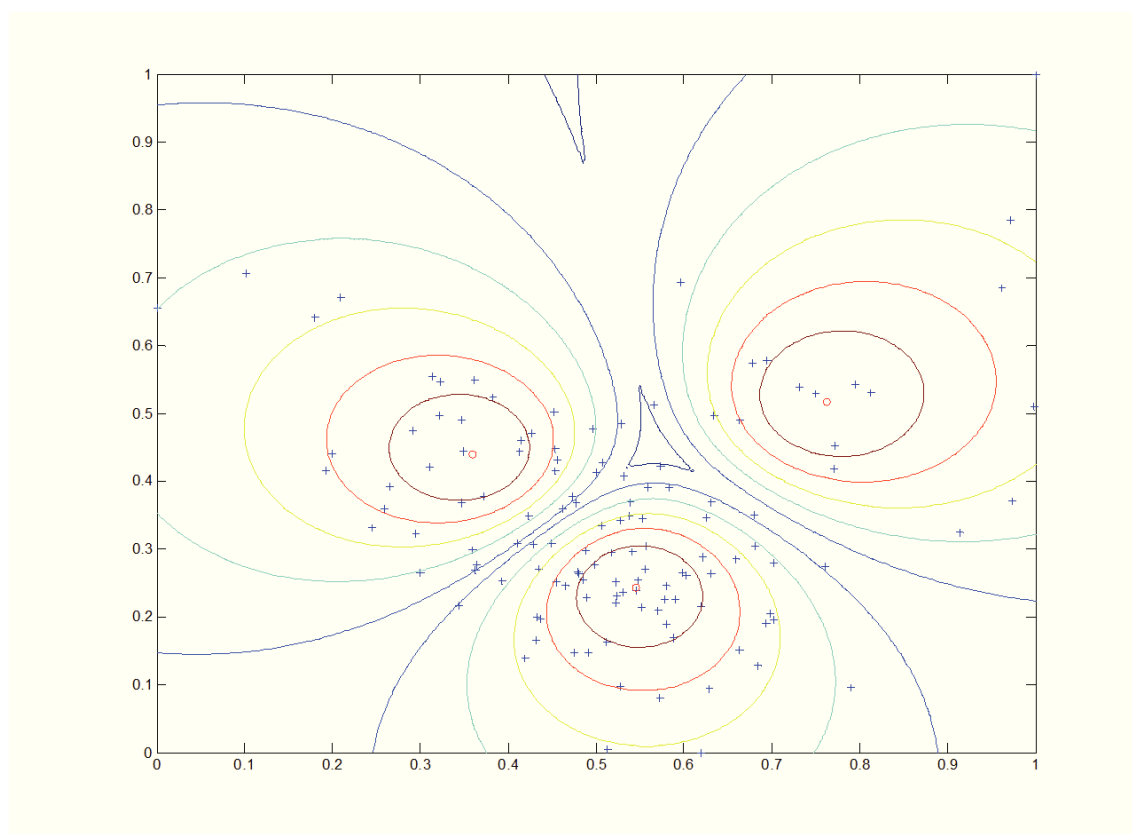


Obrázek 7.1: Klasifikace v testu A

do rozmezí hodnot 0 až 1.

Obrázek 7.2 zobrazuje klasifikaci v testu B, opět provedeného pomocí dvou hlavních komponent. Data jsou normalizována algoritmem range a klasifikována. Test C byl proveden obdobným způsobem, ale na prvních 4 hlavních komponentách vstupních parametrů. To má za následek, že nelze jednotlivé shluky graficky reprezentovat.

Výsledkem klasifikace je umístění shluků v prostoru, jejich struktura a umístění centroidů. Toto bude dále použito v návrhu vyhodnocovacího algoritmu a testování.



Obrázek 7.2: Klasifikace v testu B

Kapitola 8

Návrh vyhodnocovacího algoritmu

Po návrhu normalizace, analýzy hlavních komponent a klasifikačního algoritmu je nutné přistoupit k jádru celé práce. Tím je vytvoření algoritmu, který dokáže do stanoveného klasifikačního prostoru zařadit dalšího pacienta a zjistit jeho příslušnosti k navrženým shlukům. Tím lze později vyvodit jeho diagnózu ze souhrnných vlastností skupiny.

8.1 Normalizace nových pacientů

První záležitostí vyhodnocovacího algoritmu je normalizování parametrů pacienta stejně, jak to bylo provedeno v klasifikačním algoritmu. K tomu musí být použito modifikovaných normalizačních vzorců.

$$X_2 = \frac{X_{2old} - X_{1min}}{X_{1max} - X_{1min}} \quad (8.1)$$

Algoritmus pro modifikovanou normalizaci algoritmem range je popsán vzorcem 8.1. Hodnota nového pacienta X_2 je normalizována pomocí statistických hodnot X_{1min} a X_{1max} získaných z původního statistického souboru použitého v klasifikačním modelu. Tato normalizace se využívá před klasifikací a její nevýhoda může spočívat v zařazení prvku mimo stanovený prostor v případě, že nová hodnota X_2 není v rozsahu původních hodnot X_{1min} a X_{1max} .

$$X_2 = \frac{X_{2old} - \bar{X}_1}{\sigma_{1x}} \quad (8.2)$$

Modifikovaná normalizace algoritmem var lze vypočítat vzorcem 8.2. Nová hodnota parametru pacienta X_2 je normalizována statistickými parametry původního souboru parametru. Jedná se konkrétně o střední hodnotu \bar{X}_1 a hodnotu variance σ_{1x} . Normalizace algoritmem var je potřebná k objektivizaci dat před lineární transformací pomocí matic analýzy hlavních komponent.

8.2 Koeficienty analýzy hlavních komponent

Analýza hlavních komponent byla provedena na části dat pro analýzu. Před analýzou byla data redukována a normována algoritmem var, analýza hlavních komponent vytvořila nové proměnné, které reprezentují předchozí data v určité informační hladině.

Transformace se provedla pomocí lineární transformační matice koeficientů analýzy, protože se jedná o lineární metodu.

$$\mathbf{COEFF} = \begin{pmatrix} k_{11} & k_{12} & \dots & k_{1n} \\ k_{21} & k_{22} & \dots & k_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ k_{m1} & k_{m2} & \dots & k_{mn} \end{pmatrix} \quad (8.3)$$

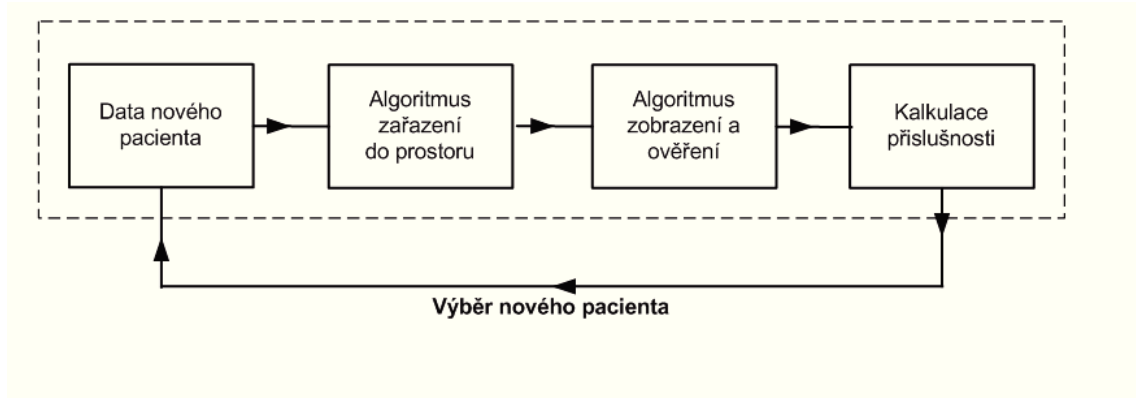
Popis této matice je ve vzorci 8.3. Matice má m řádků a n sloupců se stejným počtem, tedy matice je čtvercová a jednotlivé sloupce matice vyjadřují koeficienty jedné hlavní komponenty. To znamená, že rozměr m matice **COEFF** se shoduje s počtem parametrů pacienta, které jsou v analýze.

$$\mathbf{T}_2(p, n) = \mathbf{X}_2(p, m) \cdot \mathbf{COEFF}(m, n) \quad (8.4)$$

Rekonstrukce hlavních komponent pro nově klasifikované pacienty probíhá dle vzorce 8.4 pro maticový součin. Původní data $\mathbf{X}_2(p, m)$ jsou násobena maticí koeficientů $\mathbf{COEFF}(m, n)$, kde je volen příslušný počet sloupců n podle počtu rekonstruovaných hlavních komponent. Výsledkem je matice $\mathbf{T}_2(p, n)$, kde rozměr p značí počet nově klasifikovaných pacientů a rozměr n počet rekonstruovaných hlavních komponent. Samozřejmou podmínkou transformace je, že noví pacienti budou mít stejný počet a strukturu parametrů vstupujících do transformace.

8.3 Zařazení do klasifikačního prostoru

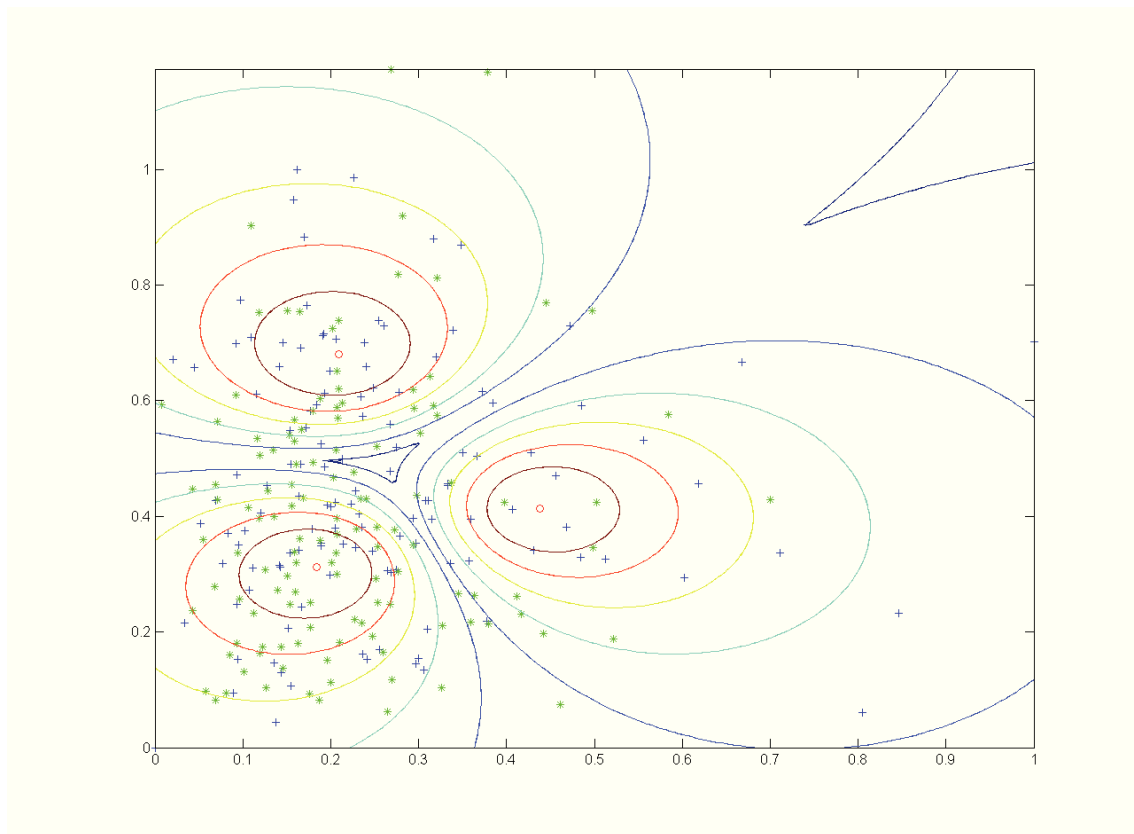
Poslední fází algoritmu rekonstrukce je zařazení nového pacienta do příslušného klasifikačního prostoru, který byl navrhnut v analýze.



Obrázek 8.1: Algoritmus zařazení do klasifikačního prostoru

Na obrázku 8.1 je zobrazen algoritmus zařazení do příslušného klasifikačního prostoru. V první fázi jsou načítána data nového pacienta po transformaci, to znamená normalizovaná modifikovaným algoritmem a pak transformovaná maticí hlavních komponent z analytické části.

V druhé části následuje samotné zařazení do klasifikačního prostoru, respektive přímo do jednotlivých klasifikačních shluků. Následuje zobrazení nového pacienta v klasifikačním prostoru a ověření správného zařazení pomocí kontrolního mechanismu.

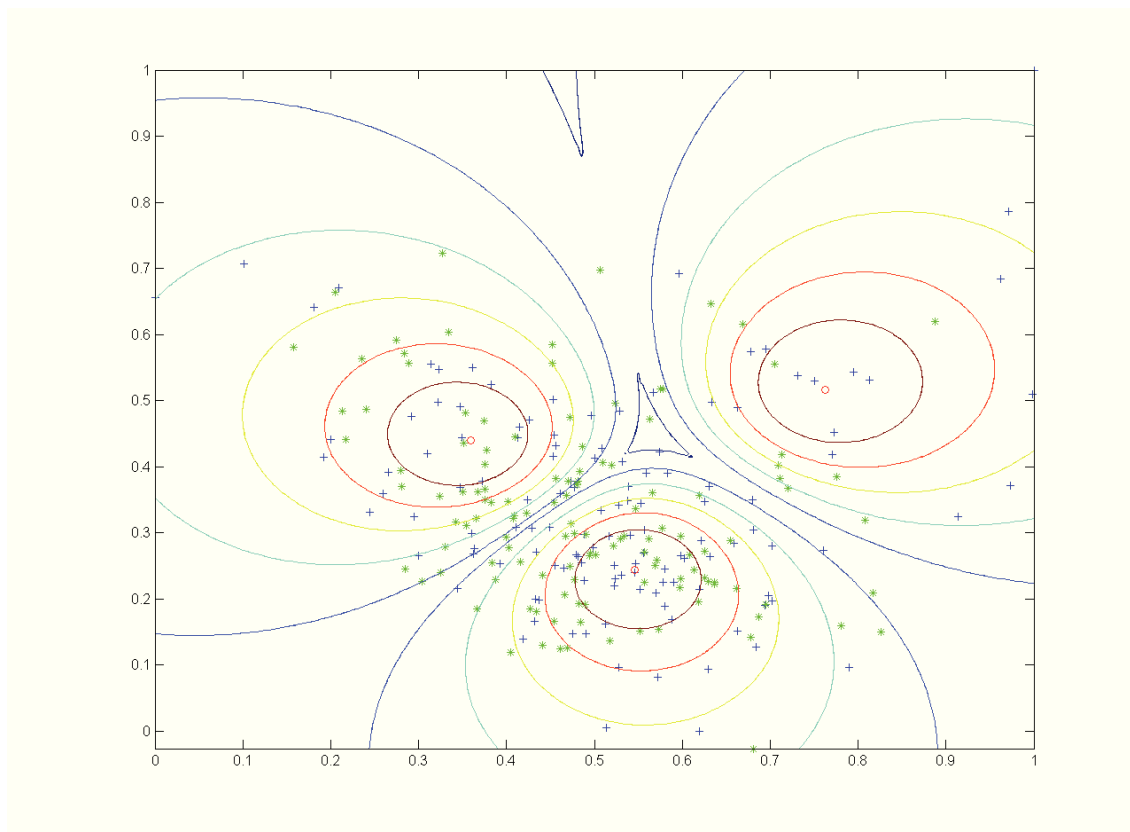


Obrázek 8.2: Zařazení do klasifikačního prostoru v testu A

Obrázek 8.2 zobrazuje zařazení do klasifikačního prostoru u testu A. Na obrázku jsou viditelné 3 shluky včetně jednotlivých vrstev a jejich centroidy znázorněné červenými body. V diagramu jsou i analytické hodnoty pacientů, které jsou zobrazeny modře. Zeleně jsou zobrazeny nově zařazené hodnoty pacientů pro testování. V testu se nacházejí dvě odlehlé hodnoty, které modifikovaná normalizace range vypočetla mimo základní rozsah 0 až 1. To může být způsobeno chybou daného měření, nebo výraznou patologií daného pacienta.

Stejným způsobem bylo provedeno zařazení do klasifikačního prostoru u testu B, které je zobrazeno na obrázku 8.3. Opět jsou zde zobrazeny 3 shluky včetně vrstev a jejich centroidy červenou barvou. Do prostoru jsou nejprve zařazené analytické hodnoty, které jsou zobrazeny modře. Potom jsou dle transformovaných hlavních komponent zařazené do prostoru testovací hodnoty pacientů, které jsou znázorněny zeleně. Žádná z nově zařazených hodnot se nenachází mimo prostor ohraničený normalizací algoritmem range.

V poslední fázi bylo provedeno zařazení do klasifikačního prostoru pro test C. Zařazení se provádí pomocí 4 hlavních komponent, což znamená, že index matice koeficientů je $n = 4$. Postup zařazení je stejný jako u testu A i testu B s tím rozdílem, že data nelze zobrazit a k dispozici jsou pouze informace o poloze daného pacienta ve čtyřrozměrném prostoru, jeho



Obrázek 8.3: Zařazení do klasifikačního prostoru v testu B

vzdálenost od centroidů jednotlivých shluků a míry příslušnosti k daným shlukům. Mimo stanovený prostor nebyli klasifikováni žádní pacienti, takže všechna data byla korektně upravena algoritmem normalizace range.

Kapitola 9

Testování navrženého algoritmu

Testování navrženého algoritmu je poslední částí práce. Korektní testování umožní ověření implementovaného postupu a zhodnocení dosažených výsledků dané klasifikace a dodatečného zařazení do klasifikačního prostoru.

9.1 Koncepce testování

Celý koncept testování je založen na ověření, zda analytická skupina pacientů s danou příslušností ke shluku má podobné statistické vlastnosti jako příslušná testovací skupina.

Testování lze rozdělit do několika částí:

1. Stanovení hranice příslušnosti
2. Testování statistických hypotéz
3. Vyhodnocení testování

V první části je nutno vybrat pacienty dle hranice příslušnosti ke shluku jak ze skupiny pro analýzu, tak i z testovací skupiny. Pak následuje otestování hypotéz pomocí výběrových testů a nakonec je provedeno souhrnné vyhodnocení testování.

9.2 Stanovení hranice příslušnosti

V algoritmu je použita fuzzy klasifikace, takže každý klasifikovaný pacient je charakterizován svou příslušností ke každému shluku a vzdáleností od jednotlivých centroidů. V principu lze tedy použít dvě cesty kalkulace příslušnosti.

1. Stanovení hranice dle vzdálenosti od jednotlivých centroidů
2. Stanovení hranice dle příslušnosti ke shluku

První postup je použitelný v případě, kdy jsou shluky rovnoměrně rozloženy, a může se použít i v klasifikacích, které nepoužívají fuzzy algoritmus přiřazení.

Druhý postup je pro navržený algoritmus vhodnější, protože reflektuje vypočítanou strukturu shluků. Každý pacient má tedy tolik příslušností, kolik byl stanoven počet shluků v klasifikaci. Důležité je tedy stanovit hranici, při které je pacient považován za významného

reprezentanta skupiny souhrnných vlastností.

$$\varphi = \frac{200}{n+1} \quad (9.1)$$

Pro stanovení hranice příslušnosti byl navržen výpočet dle vzorce 9.1. Příslušnost je vyjádřena koeficientem φ v procentech, který určuje hranici příslušnosti. Tento koeficient je nelineárně závislý na počtu shluků. Pro jeden shluk je hranice logicky 100 % příslušnosti a pro více shluků nelineárně klesá. Hranice pro 2 shluky je 66,6 %, a pro 3 shluky 50 %. Algoritmus výpočtu hranice lze použít do omezeného počtu shluků, je vypovídající zhruba do počtu 5 shluků, pak je již hranice příslušnosti příliš nízká a zanikají rozdíly mezi shluky.

Ve všech analýzách byla použita klasifikace do 3 shluků a hraniční míra klasifikace tedy bude použita 50 %. Mohou tedy nastat 2 případy:

1. Hodnoty příslušnosti se u všech shluků pohybují do hodnoty 50 %.
2. Jedna hodnota příslušnosti překročí 50 %.

V prvním případě se jedná o pacienta, který má rovnoměrně vlastnosti více shluků, tento pacient je tedy z vyhodnocení vyloučen. Testování je založeno na tom, že se pacient projeví jako většinový (více než 50 % příslušnosti) zástupce daného shluku.

9.3 Testování statistických hypotéz

V další fázi je nutno porovnat právě dva statistické soubory každého shluku, a to:

1. Soubor parametrů pacienta pro analýzu překračující hranici příslušnosti k danému shluku
2. Soubor parametrů pacienta pro testování překračující hranici příslušnosti k danému shluku

K porovnávání jsou použity dvě základní metody testování statistických hypotéz, a to T-test a F-test. Oba testy se používají k testování statistických souborů, u kterých můžeme předpokládat normální rozdělení [4].

K ověření algoritmu se musí porovnat statistický soubor pacientů jednotlivých shluků určený hranicí příslušnosti. T-test testuje hypotézu, zda dvě normální rozdělení mající stejný rozptyl pocházejí ze stejného statistického souboru, pomocí střední hodnoty, čili mají stejné statistické vlastnosti. Pro kontrolu, zda mají oba statistické soubory stejný rozptyl, je použit F-test [4].

Testování bylo provedeno v prostřední MATLAB pomocí předdefinovaných funkcí `ttest2` a `vartest2`, které provádějí dvouvýběrové testy. Takto lze ověřit hypotézu, zda oba statistické soubory mají stejné statistické vlastnosti.

9.4 Vyhodnocení testování

Celkově bylo provedeno vyhodnocení testování třemi způsoby:

1. Grafická evaluace

2. Výpis statistických parametrů

3. Testování hypotéz

V příloze jsou uvedeny krabicové diagramy testování, diagramy parametrů testu A jsou zobrazeny na obrázcích D.1, D.2 a D.3. Test B je popsán obrázky D.4, D.5, D.6 a D.7. Poslední test C je znázorněn na obrázcích D.8, D.9, D.10, D.11, D.12 a D.13.

V grafu je znázorněn dvojnásobný počet statistických parametrů jednoho souboru, oproti celkovému počtu shluků v klasifikaci. Jsou zobrazeny vždy dvojice statistických souborů, a to skupina pro analýzu a testovací skupina. Krabicový diagram 1 odpovídá analytické skupině shluku A, diagram 2 testovací skupině shluku A. Diagramy 3 a 4 náleží shluku B, diagramy 5 a 6 shluku C. Graficky lze takto porovnat, zda má testovací skupina stejné statistické vlastnosti, jako skupina pro analýzu.

Skupiny pacientů lze porovnat pomocí přímého srovnání statistických parametrů. První vlastností k porovnání je medián \tilde{x} , což je hodnota, která dělí řadu parametrů seřazených dle velikosti na dvě stejné poloviny. V případě spojitého rozdělení zadaného hustotou pravděpodobnosti f platí vzorec 9.2 [3].

$$\int_{-\infty}^{\tilde{x}} f(x)dx = 0,5 \quad (9.2)$$

Dalším parametrem, který může být použit pro srovnání, je střední hodnota statistického souboru \bar{x} . Střední hodnota je definována jako vážený průměr daného rozdělení vyjádřena vzorcem 9.3 pro diskrétní náhodnou veličinu s rozdělením pravděpodobností $p_x(x_i)$, kde i je množina hodnot statistického souboru [3].

$$\bar{x} = \sum_i x_i \cdot P(x_i) \quad (9.3)$$

Poslední hodnotou pro porovnávání je rozptyl, nebo také střední kvadratická odchylka σ^2 . Rozptyl je definovaný jako střední hodnota kvadrátů odchylek od střední hodnoty \bar{x} . Rozptyl je definován vzorcem 9.4 pro diskrétní statistický soubor. [3].

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (9.4)$$

Takto vypočtené hodnoty v příloze pro test A jsou uvedené v tabulce E.3. Pro test B jsou vypočteny v tabulkách E.8 a E.9. Závěrečný test C je popsán údaji v tabulkách E.16, E.17 a E.18. Hodnoty v tabulkách označené číslem 1 se týkají analytické skupiny a hodnoty označené číslem 2 se týkají testovací skupiny. Písmena A, B a C označují jednotlivé shluky, ve sloupcích tabulky jsou uvedeny názvy jednotlivých parametrů.

Takto navržené porovnávání je velice náročné, proto byla zvolena evaluace pomocí testování hypotéz, konkrétně T-test a F-test. Testování bylo rozděleno do dvou částí:

1. Souhlasný test – evaluace správného zařazení
2. Křížový test – evaluace dobré separace

Souhlasný test ověřuje základní vlastnost algoritmu zařazení do klasifikačního prostoru a to, jestli byli zařazení pacienti se stejnými statistickými parametry do stejného shluku.

Křížový test ověřuje funkci prvotního návrhu klasifikace, a to jestli byla navržena taková klasifikace, která jasně odděluje jednotlivé statistické skupiny v daném parametru.

Výsledky testování statistických hypotéz se nacházejí v příloze, pro test A je to tabulka E.1 pro souhlasný test a tabulka E.2 pro křížový test. V testu B to jsou tabulky E.4 a E.5 pro souhlasný test, tabulky E.6 a E.7 pro křížový test. Pro poslední test C jsou to tabulky E.10, E.11 a E.12 pro souhlasný test, dále tabulky E.13, E.14 a E.15 pro křížový test.

V tabulkách jsou písmenem F označeny výsledky F-testu a písmenem T výsledky T-testu. V případě, že je hypotéza přijata a statistické soubory jsou považovány, že pochází z jednoho statistického souboru na hladině významnosti 0,05, výsledek je potom označen jako 0. V případě, že je hypotéza odmítnuta, je hodnota výsledku testu 1. Rovněž je pod každým výsledkem testování hypotézy vypočtena p hodnota. Za lepší výsledek lze považovat to, když se u souhlasného testu maximální počet nulových hypotéz přijme a u křížového testu odmítne. Vypovídající hodnotou je potom výsledek T-testu, F-test slouží jako podklad, jestli je T-test prováděn korektně.

$$\epsilon_s = 1 - \frac{\sum_{j=1}^c \sum_{i=1}^n H_{ij}}{n \cdot c} \quad (9.5)$$

Pro celkové vyhodnocení souhlasného T-testu u jednotlivých návrhů byl navržen vztah 9.5. Koeficient souhlasného testu ϵ_s vyjadřuje míru, jak dobře funguje algoritmus zařazení do klasifikačního prostoru, kde 1 znamená maximální hodnotu nejlepšího zařazení. Koeficient se počítá z testování všech parametrů n a navržených shluků c . Hodnota H je výsledek testování hypotézy, kde 0 znamená přijetí nulové hypotézy a 1 odmítnutí nulové hypotézy, to vše na hladině významnosti 0,05.

$$\epsilon_k = \frac{\sum_{j=1}^c \sum_{i=1}^n H_{ij}}{n \cdot c} \quad (9.6)$$

Vyhodnocení křížového T-testu se provádí dle vzorce 9.6. Koeficient křížového testu ϵ_k vyjadřuje, jak moc jsou izolovány jednotlivé shluky ve všech parametrech. Hodnota 1 znamená, že klasifikovaný pacient se unikátně zařadí v každém parametru do příslušného shluku. Parametr se počítá pomocí všech údajů pacienta n a navržených shluků c . Hodnota H je určena stejně, jako u výpočtu souhlasného koeficientu.

Vyhodnocení F-testu se pro souhlasný i křížový test provádí dle navrženého vzorce 9.5.

Tabulka 9.1: Evaluace navržených testů

Test	F-test souhlasný	T-test souhlasný	F-test křížový	T-test křížový
Test A	76,6 %	100 %	60,0 %	70,0 %
Test B	70,8 %	89,6 %	68,8 %	66,7 %
Test C	87,0 %	97,8 %	65,2 %	60,9 %

Tabulka 9.1 popisuje souhrnné výsledky souhlasného a křížového testu statistických hypotéz. Souhlasný F-test vykazuje nejlepší výsledek 87,0 % u testu C a nejhorší výsledek 70,8 % u testu B, tedy nejpřesněji je provedeno vyhodnocení u testu C, poté u testu A a nakonec až u testu B v souhlasném T-testu. T-test souhlasný vykazuje nejlepší hodnotu u testu A, a to rovných 100 %, nicméně hodnota testu C 97,8 % se tomuto výsledku blíží.

U křížového testu vykazují hodnoty F-testu podobné výsledky v rozmezí 60,0 % až 68,8 %. Vzhledem k tomu, že se jedná o nesouhlasné testování, jsou výsledky odpovídající. Křížový T-test vykazuje nejlepší izolaci parametrů u testu A a to 70,0 %, nejhorší u testu C 60,9 %. S přihlédnutím k tomu, že test A obsahuje pouze 10 parametrů analýzy a test C 23 parametrů, lze považovat výsledek u testu C za výborný.

Z testování vyplývá, že nejlepší výsledky prokazuje algoritmus navržený pomocí testu C. V souhlasném testu vykazuje největší přesnost testování pomocí F-testu a T-test prokazuje téměř dokonalou přesnost 97,8 %. Izolační hodnota je na velmi dobré úrovni, to znamená, že 60,9 % parametrů testu vykazuje specifické hodnoty pro každý shluk. Test A se umístil v hodnocení jako druhý, protože sice má 100 % úspěšnost v souhlasném T-testu, ale předpoklad správné variace ověřovaný F-test již tak úspěšný není. Hodnota izolace je 70 %, ale pouze na 10 parametrech testu. Nejhorší ze srovnání vychází test B, který má nejhorší výsledky souhlasného T-testu i F-testu. Izolace je ve standardních hodnotách testu A i C. To může být způsobeno špatnou aproximací shluků, kdy nešlo jednoznačně určit správný počet shluků dle použitých indikátorů.

V praxi lze tedy doporučit použití algoritmu navrženého testem C, protože využívá nejvíce parametrů a klasifikace probíhá pomocí čtyřrozměrného prostoru. S výhodou lze použít i algoritmus navržený testem A, ovšem jen pro malý počet parametrů.

9.5 Medicínské zhodnocení shluků

Souhrnné hodnocení jednotlivých parametrů bylo provedeno u testu C. Byly vybrány vlastnosti specifické vždy pro jednotlivý shluk, hodnocení bylo provedeno ve spolupráci s lékařem.

V tabulce 9.2 jsou uvedeny průměrné hodnoty jednotlivých parametrů shluku A. Průměrný pacient v tomto shluku trpí obezitou II. stupně a po podání glukózy se velmi omezí jeho schopnost spalovat tuky. To vyplývá z hodnot přístroje DeltaTrac II, kdy spalování fat1 na lačno je 42 %, ale spalování fat2 a fat3 je 17,53 %, respektive 16 %. Průměrné hodnoty inzulínu po podání glukózy jsou rovněž na velmi vysoké úrovni. Průměrný pacient obsahuje pouze 53,87 % aktivní tělesné hmoty a 45,22 % tukové tkáně. Celkový součet aktivní tělesné hmoty a tukové tkáně je 100 %, odchylka 0,91 % je způsobena zaokrouhlováním na celá čísla při měření. Hodnoty cholesterolu jsou zvýšeny na 5,31.

Tabulka 9.3 popisuje průměrného pacienta, který byl zařazen do shluku B. Tento pacient také trpí obezitou II. stupně, stejně jako ve shluku A, nicméně má lepší spalování tuku, a to v parametru fat1 51 %. Parametry fat2 a fat3 po podání glukózy mají hodnoty 40,66 %, respektive 31 %. Pacient má rovněž zvýšené parametry inzulínu ins2 a ins3, nicméně parametr ins3 má nižší úroveň a dochází tedy rychleji ke snižování inzulínu v krvi. S tím korespondují parametry aktivní tělesné hmoty a tukové tkáně, aktivní tělesná hmota tvoří 66,35 % celkové hmotnosti pacienta a tuková tkáň 32,65 % těla. Cholesterol je zde v normě.

Průměrný pacient shluku C je popsán tabulkou 9.4. Pacient má pouze nadváhu, ale jeho spalování tuku je v úrovních pacienta ve shluku A. Inzulín je na nižších hodnotách než u pacientů obou předchozích shluků. Kompozice těla je podobná jako u pacienta shluku B, což dokazují hodnoty aktivní tělesné hmotnosti 64,56 % a tukové tkáně 34,59 %. Cholesterol je zde zvýšený.

Z celkového hodnocení vyplývá, že v jednotlivých shlucích se umístily tyto typy pacientů:

- Shluk A - pacient s obezitou, který není schopen odbourávat tuky, jeho tělo tvoří téměř z poloviny tuk a hodnoty inzulínu jsou vysoké.

Tabulka 9.2: Medicínské zhodnocení, shluk A

Parametr	Hodnota	Zhodnocení	Specifická hodnota
BMI	37,6	Obezita II. stupně	ANO
fat1	42	—	ANO
fat2	17,53	—	ANO
fat3	16	—	ANO
ins1	70	V normě	NE
c-p1	0,7	Fyziologický	NE
ins2	551	Velmi zvýšený	ANO
c-p2	2,96	Fyziologický	NE
ins3	423	Velmi zvýšený	ANO
c-p3	2,78	Fyziologický	NE
tuk	45,22	—	ANO
ath	53,87	—	ANO
voda	38,91	—	NE
bezv	14,87	—	NE
ect	17,43	—	NE
ctv	40,74	—	NE
alt	0,53	V normě	NE
moč	312	V normě	NE
ldlv	3,23	V normě	NE
hdlc	1,53	V normě	NE
chol	5,31	Zvýšený	ANO
glu1	8,76	V normě	NE
glu2	6,96	V normě	NE

- Shluk B - pacient s obezitou, který odbourává tuky, jeho tělo tvoří z jedné třetiny tuk.
- Shluk C - pacient s nadváhou, který má problém odbourávat tuky, nicméně jeho tělesná hmotnost není vysoká a tuk tvoří jednu třetinu jeho těla.

Námětem k dalšímu zkoumání by mohla být analýza těchto skupina z hlediska časového průběhu odbourávání tuku, nebo časového průběhu insulínu a c-peptidu. Je možno zkoumat i jednotlivé parametry krevního rozboru a měření impedance tkáně, popřípadě důvodu obezity jednotlivých typů. Toto však není cílem této práce, ale dalšího medicínského výzkumu, pro který můžou sloužit výsledky této práce jako podklad.

Tabulka 9.3: Medicínské zhodnocení, shluk B

Parametr	Hodnota	Zhodnocení	Specifická hodnota
BMI	37,4	Obezita II. stupně	ANO
fat1	51	—	ANO
fat2	40,66	—	ANO
fat3	31	—	ANO
ins1	80	V normě	NE
c-p1	0,97	Fyziologický	NE
ins2	557	Velmi zvýšený	ANO
c-p2	3,37	Fyziologický	NE
ins3	341	Zvýšený	ANO
c-p3	2,48	Fyziologický	NE
tuk	32,65	—	ANO
ath	66,35	—	ANO
voda	47,90	—	NE
bezv	19,85	—	NE
ect	18,90	—	NE
ctv	48,10	—	NE
alt	0,72	V normě	NE
moč	366	V normě	NE
ldlv	2,78	V normě	NE
hdlc	1,26	V normě	NE
chol	4,49	V normě	NE
glu1	8,02	V normě	NE
glu2	6,14	V normě	NE

Tabulka 9.4: Medicínské zhodnocení, shluk C

Parametr	Hodnota	Zhodnocení	Specifická hodnota
BMI	28,1	Nadváha	ANO
fat1	30	–	ANO
fat2	12	–	ANO
fat3	10	–	ANO
ins1	46	V normě	NE
c-p1	0,56	Fyziologický	NE
ins2	331	Zvýšený	ANO
c-p2	2,42	Fyziologický	NE
ins3	270	Zvýšený	ANO
c-p3	2,42	Fyziologický	NE
tuk	34,59	–	ANO
ath	64,56	–	ANO
voda	47,94	–	NE
bezv	11,94	–	NE
ect	22	–	NE
ctv	49,41	–	NE
alt	0,43	V normě	NE
moč	253	V normě	NE
ldlv	3,53	V normě	NE
hdlc	1,91	V normě	NE
chol	5,86	Zvýšený	NE
glu1	8,56	V normě	NE
glu2	6,89	V normě	NE

Kapitola 10

Závěr

10.1 Zhodnocení práce

V první části této práce byly popsány teoretické znalosti v oblasti metabolismu a energometrických testů. Důležitou součástí je měření respiračního kvocientu a utilizace jednotlivých složek potravy během několika hodinových testů. V technickém rozboru řešení byly popsány základní metody získávání dat pro účely diagnostiky metabolismu a metody pro popis organismu pacienta. Z předložených typů diagnostických testů byl vybrán typ 1, který zahrnuje kompletní diagnostiku všech parametrů pacienta a probíhá ve třech etapách po celkovou dobu tří hodin.

Práce využívá tři základních metod pro získání dat. První data respiračního kvocientu jsou získávána z přístroje DeltaTrac II, kde jsou analyzovány respirační plyny, a z nich vypočítány parametry utilizace složek potravy. Dále je zde vytvářen základní popis pacienta. Druhým zdrojem dat jsou biochemické testy, v první řadě to je základní rozbor krve pacienta a poté i hodnoty inzulínu a c-peptidu. Posledním zdrojem dat je přístroj pro měření impedance tkáně QuadScan 4000, který zjišťuje strukturu lidského těla, hodnoty rozložení vody v organismu a doplňující popis pacienta.

V praktické části byly provedeny prvotní statistické testy, které popsaly soubor pacientů pomocí standardních statistických veličin a graficky jej zobrazily. Byly navrženy tři testovací postupy a v každém postupu byla provedena analýza hlavních komponent, pomocí které byla navržena redukce parametrů jednotlivých testů, byla zjištěna informační míra jednotlivých parametrů a hlavních komponent. Tato redukce byla zhodnocena a popsána v příloze této práce.

Dále byla navržena metodika fuzzy klasifikace redukováných dat a použita normalizace dle dvou základních algoritmů. Byly použity základní ukazatele počtu klasifikačních skupin a pomocí nich stanoveny počty shluků pro jednotlivé testy. Samotný algoritmus zařazení pacientů do klasifikačního prostoru byl navržen pomocí lineární transformace maticí hlavních komponent, modifikované normalizace a fuzzy klasifikace. Výsledkem bylo zařazení analytické skupiny pacientů do jednotlivých shluků dle klasifikačních pravidel.

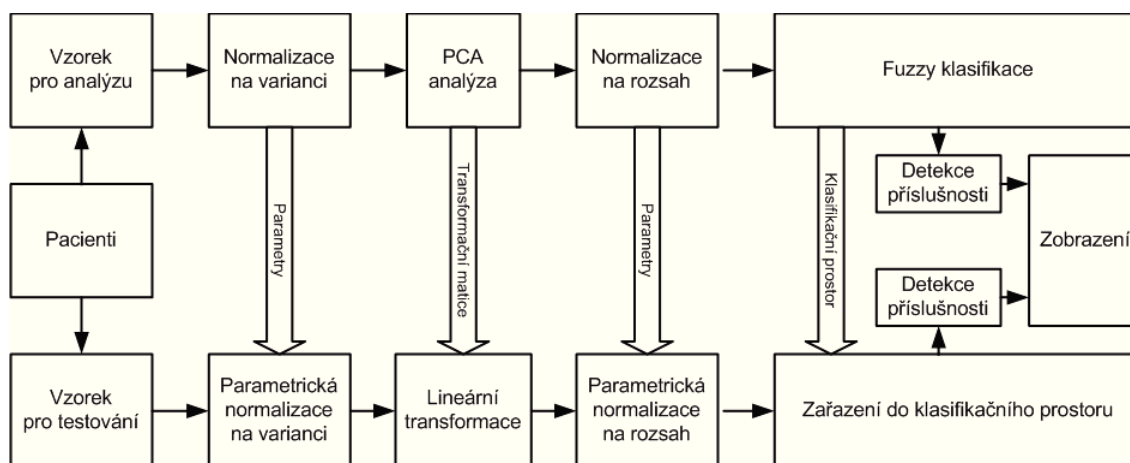
Testování proběhlo pomocí stanovení příslušnosti ke shluku a standardního testování statistických hypotéz, T-testu a F-testu. K vyhodnocení pak byl použit vlastní algoritmus a všechny testy byly vyhodnoceny pomocí souhlasného a křížového algoritmu. Pomocí tohoto testu byl stanoven nejlepší testovací algoritmus a ověřena separační vlastnost algoritmu.

Výsledkem práce je algoritmus pro navržení klasifikačního prostoru a zařazení do něj.

Pro data použitá v práci je nejvhodnější testovací postup C, který využívá klasifikaci do čtyřrozměrného prostoru pomocí 23 základních parametrů pacienta. Algoritmus C má vynikající výsledky souhlasného testu a zároveň dobře separuje vlastnosti parametrů v jednotlivých shlucích. Tento algoritmus se dá využít v mnoha aplikacích popsaných v práci, zejména však jako automatizovaný diagnostický systém v metabolických poradnách a odborných pracovištích. Algoritmus lze jednoduše implementovat, což je podstatná výhoda proti algoritmu v předběžné studii [18].

10.2 Rekapitulace postupu řešení

Pro řešení práce byl navrhnut celkový algoritmus postupu, tento algoritmus je schématicky znázorněn na obrázku 10.1. Postup popisuje cestu dat zpracováním, v první fázi jsou pacientská data rovnoměrně rozdělena na vzorek pro analýzu a vzorek pro testování.



Obrázek 10.1: Navržený postup řešení práce

Skupina pro analýzu slouží k navržení algoritmu pro klasifikaci metabolických parametrů. Surová data se nejprve normalizují dle variance daného parametru, a pak se provede analýza hlavních komponent, která odhalí vazby jednotlivých parametrů a umožní redukovat počet parametrů k následné klasifikaci.

Před samotnou klasifikací se provede normalizace dle rozsahu, aby bylo docíleno korektního zobrazení v klasifikačním algoritmu. Pak následuje klasifikace s fuzzy algoritmem, neboť tento postup umožní zařadit pacienty do jednotlivých shluků a zjistit i míry příslušnosti jednotlivých shluků. Na konci řetězce je nastavitelná detekce příslušnosti k jednotlivým shlukům a zobrazení výsledků.

Vzorky pacientů pro testování procházejí obdobným řetězcem s výjimkou toho, že parametry k jednotlivým úkonům jsou přebírány z analytického řetězce. Prvně se provede parametrická normalizace dle variance zjištěné v analýze, pak následuje lineární transformace dle zjištěných transformačních matic analýzy hlavních komponent.

Před klasifikací se opět provede normalizace na rozsah dle parametrů zjištěných v analýze a následuje zařazení do klasifikačního prostoru. Algoritmus postupně zařazuje testovací pacienty do stávajícího klasifikačního prostoru. Pak následuje nastavitelná detekce příslušnosti

ke shlukům a zobrazení. Způsob tohoto zobrazení umožňuje porovnat statistické parametry analytického vzorku a vzorku testovacího.

10.3 Praktické uplatnění práce

V práci byly navrženy postupy, které lze využít v primárním lékařském výzkumu. Pomocí algoritmu analýzy hlavních komponent lze odhalit metabolické a jiné vazby mezi jednotlivými parametry. Tímto lze odhalit provázanost jednotlivých procesů v lidském organismu a dále stanovit jejich informační důležitost pro diagnostiku. Pomocí navržených postupů lze zjistit diagnostický význam jednotlivých parametrů a citlivost diagnostiky po odstranění některých hodnot. Dále lze vyloučit prvky, které zanáší chybu diagnostiky a zkvalitnit správné zařazení pacienta do léčby.

Druhou aplikační oblastí je hodnocení ekonomičnosti a efektivity vyšetření pacienta. Pomocí algoritmu redukce parametrů a hodnocení výsledku, lze evaluovat, jaké parametry přinášejí důležité informace do diagnostiky. Pomocí tohoto algoritmu lze vyšetření zjednodušit a zlepšit jeho efektivitu. Toto ve finále zlepší ekonomičnost a dostupnost vyšetření širší oblasti pacientů.

Poslední a nejdůležitější oblastí je aplikace automatizovaných diagnostických systémů. Navržený systém klasifikace a možnost dodatečného zařazení do klasifikačního prostoru umožňuje navrhnout přesný automatizovaný systém hodnocení pacienta. Při stanovení charakteristických souhrnných vlastností shluků lze po zařazení do klasifikačního prostoru automaticky zjistit diagnózu pacienta.

Navržený systém automatizované diagnostiky lze zhodnotit pomocí navržené metodiky testování a určit přesnost zařazení pacienta do příslušného shluku. Při návrhu mnohoslukového uspořádání lze diagnostikovat jednotlivé symptomy pacienta. Při omezení klasifikace na malý počet shluků je diagnostika robustní a umožňuje zařazení pacienta do jednoho z velkých shluků.

Algoritmus lze kromě lékařství využít i v průmyslových odvětvích k hodnocení vlastností produktů, nebo také k detekci souhrnných vlastností měřených objektů. Aplikovat rovněž lze k automatickému zjišťování typu objektu dle definovaných vzorů.

10.4 Doporučení pro další postup

Další postup je důležitý hlavně v oblasti praktického nasazení algoritmu jako diagnostického nástroje. Je potřeba stanovit referenční skupiny pacientů pro určení vzoru diagnostiky. U těchto skupin je nutné stanovit soubor souhrnných vlastností a míru příslušnosti k těmto vlastnostem. Po tomto návrhu je potřeba vytvořit kontrolní lékařské testování, kdy budou testováni pacienti s jasnou diagnózou a po automatickém zařazení budou diagnózy porovnány.

Z hlediska aplikačního pohledu je nutné najít vhodnou platformu pro implementaci algoritmu. Může se jednat o základní programovací prostředky nebo finální rozhodovací zařízení na úrovni hardwarového návrhu. Navržený systém není výpočetně náročný, při současném výkonu počítačů trvala simulace v programu MATLAB v řádu sekund. V případě speciálně navrženého prostředí pro tuto aplikaci bude diagnostická doba menší než jedna sekunda.

Algoritmus může být rovněž zachován v prostředí MATLAB a upraven pro další možnosti primárního lékařského výzkumu v oblasti zkoumání metabolických dějů.

Literatura

- [1] ABONYI, János; FEIL, Balázs. *Cluster Analysis for Data Mining and System Identification*. Berlin : Birkhäuser Verlag AG, 2007. 303 s. ISBN 978-3-7643-7987-2.
- [2] ALMEIDA, R.J.; SOUSA, J.M.C.; Comparison of fuzzy clustering algorithms for classification. In *Evolving Fuzzy Systems, 2006 International Symposium on* , vol., no., pp.112-117, Sept. 2006
- [3] ANDĚL, Jiří. *Matematika náhody*. Třetí upravené vydání. Praha : MATFYZPRESS, 2007. 290 s. ISBN 80-7378-004-6.
- [4] ANDĚL, Jiří. *Statistické metody*. Druhé přepracované vydání. Praha : MATFYZPRESS, 1998. 274 s. ISBN 80-85863-27-8.
- [5] ANDRESOVÁ, Markéta; NOVÁK, Ivan. Měření energetické potřeby organismu metodou nepřímé kalorimetrie a její užití v praxi. In *Pediatric po promoci*. Ročník I, číslo 1/2004. Praha, 2004. s. 56-60.
- [6] BAILEY NJ, SAMPSON J, HYLANDS PJ, NICHOLSON JK, HOLMES E. Multi-component metabolic classification of commercial feverfew preparations via high-field 1H-NMR spectroscopy and chemometrics. In *Planta Med*. 2002 Aug;68(8):734-8. PubMed PMID: 12221598.
- [7] BALASKO, Balazs; ABONYI, Janos; FEIL, Balazs. *Fuzzy Clustering and Data Analysis Toolbox* [online]. Hungary : Department of Process Engineering at the University of Veszprem, 2005 [cit. 2011-05-29]. Dostupné z WWW: <http://www.fmt.vein.hu/softcomp/fclusttoolbox/FuzzyClusteringToolbox.pdf>.
- [8] BASS, A., BEDICZKA, D., EYER, P., HOFER, S. and PETTE, D., Metabolic Differentiation of Distinct Muscle Types at the Level of Enzymatic Organization. In *European Journal of Biochemistry*, 10: 198206. doi: 10.1111/j.1432-1033.1969.tb00674.x
- [9] BEZDEK J.C.; DUNN J.C. Optimal fuzzy partitions: A heuristic for estimating the parameters in a mixture of normal distributions. In *IEEE Transactions on Computers*, pages 835-838, 1975.
- [10] ČERMÁK, P.; CHMIEL, P.; Parameters optimization of fuzzy-neural dynamic model. In *Fuzzy Information, 2004. Processing NAFIPS '04. IEEE Annual Meeting of the* , vol.2, no., pp. 762- 767 Vol.2, 27-30 June 2004

- [11] *Development Cycles Tutorials* [online]. 2004 [cit. 2011-07-02]. Entity Relationship Modeling. Dostupné z WWW: <http://www.devarticles.com/c/a/Development-Cycles/Entity-Relationship-Modeling/>.
- [12] DU BOIS, Eugene F. Clinical Calorimetry : A Graphic Representation of the Respiratory Quotient and the Percentage of Calories from Protein, Fat, and Carbohydrate. In *The Journal of Biological Chemistry*. New York, 1923. s. 43-49.
- [13] *Deltatrac II : Weltweit der genaueste und vielseitigste Metabolic Monitor*. Finland : Datex-Ohmeda Division, 2003. 115 s.
- [14] FAUZI BIN OTHMAH, Mohd; YAU, Thomas Moh Shan; Neuro Fuzzy Classification and Detection Technique for Bioinformatics Problems. In *Modelling & Simulation, 2007. AMS '07. First Asia International Conference on* , vol., no., pp.375-380, 27-30 March 2007
- [15] HAN Jing; Application of Fuzzy Data Mining Algorithm in Performance Evaluation of Human Resource. In *Computer Science-Technology and Applications, 2009. IFCSTA '09. International Forum on*, vol.1, no., pp.343-346, 25-27 Dec. 2009
- [16] HANUŠ, Jan. *Shluková analýza a její aplikace*. Plzeň, 2009. 42 s. Bakalářská práce. Západočeská univerzita v Plzni.
- [17] HARDLE, Wolfgang; SIMAR, Leopold. *Applied Multivariate Statistical Analysis* [online]. Morgantown : West Virginia University, 2003-05-29 [cit. 2011-07-02]. Dostupné z WWW: <http://www.stat.wvu.edu/~jharner/courses/stat541/mva.pdf>.
- [18] HLAVICA, Jakub. *Klasifikace komplexních metabolických patientských dat*. Ostrava, 2011. 52 s. Bakalářská práce, vedoucí Ing. Michal Prauzek. Vysoká škola báňská – Technická univerzita Ostrava.
- [19] HLÚBIK, J.; HLÚBIK, P.; LHOTSKÁ, L. Body Impedance Analysis. In *World Congress on Medical Physics and Biomedical Engineering, September 7 - 12, 2009, Munich, Germany*. Berlin : Springer Berlin Heidelberg, 2009. s. 842-844.
- [20] HOPNER F.; KLAWONN F.; KRUSE R.; RUNKLER T. *Fuzzy Cluster Analysis*. Wiley, Chichester, 1999.
- [21] JAHANSHAHI, J.A.; GHORASHI, S.A.; SADREAZAMI, H.; Jamming detection at base station using fuzzy c-means algorithm. In *Computer Networks and Distributed Systems (CNDS), 2011 International Symposium on* , vol., no., pp.40-44, 23-24 Feb. 2011
- [22] JANČÍK, Jiří; ZÁVODNÁ, Eva; NOVOTNÁ, Martina. *Fyziologie tělesné zátěže vybrané kapitoly* [online]. Brno : Fakulta sportovních studií MU, 2006 [cit. 2011-07-01]. Dostupné z WWW: <http://is.muni.cz/elportal/estud/fsps/js07/fyzio/texty/index.html>
- [23] JIABIN Deng; JUANLI Hu; HEHUA Chi; JUEBO Wu; An Improved Fuzzy Clustering Method for Text Mining. In *Networks Security Wireless Communications and Trusted Computing (NSWCTC), 2010 Second International Conference on* , vol.1, no., pp.65-69, 24-25 April 2010

- [24] KRČ, Ivo. Hematologie – hodnocení krevního obrazu. *Urológia v skratke*. 2007, 56, s. 231-232. Dostupný také z WWW: http://www.solen.sk/index.php?page=pdf_view&pdf_id=2830
- [25] LANGMEIER, Miloš a kol. *Základy lékařské fyziologie*. 1. vyd. Praha : Grada, 2009. 320 s. ISBN 978-80-24725-26-0.
- [26] MARTINÍK, Karel. *Obezita, nadváha: Od teorie k praxi*. 1. vydání. Hradec Králové : Garamon s.r.o., 2008. 151 s. ISBN 978-80-86472-37-9.
- [27] MCLELLAN, S.; WALSH, T.; BURDESS, A. Comparison between the Datex-Ohmeda M-COVX metabolic monitor and the Deltatrac II in mechanically ventilated patients. In *Intensive Care Medicine*. 2002-07-01, 28, 7, s. 870-876.
- [28] *MySQL* [online]. 2011 [cit. 2011-07-02]. MySQL 5.6 Reference Manual. Dostupné z WWW: <http://dev.mysql.com/doc/refman/5.6/en/index.html>
- [29] NELSON, David L.; COX, Michael M. *Lehninger Principles of Biochemistry*. New York : W. H. Freeman and company, 2005. 841 s. Dostupné z WWW: <http://bcs.whfreeman.com/lehninger/default.asp>. ISBN 0-7167-4339-6.
- [30] ODUKOYA, O.H.; ADEROUNMU, G.A.; ADAGUNODO, E.R.; An Improved Data Clustering Algorithm for Mining Web Documents. In *Computational Intelligence and Software Engineering (CiSE), 2010 International Conference on* , vol., no., pp.1-8, 10-12 Dec. 2010
- [31] PAUL, R.; HOQUE, A.S.M.L.; Clustering medical data to predict the likelihood of diseases. In *Digital Information Management (ICDIM), 2010 Fifth International Conference on* , vol., no., pp.44-49, 5-8 July 2010
- [32] RICHARDSON, Henry B. Clinical Calorimetry : The Effect of the Absence of Sweat Glands on the Elimination of Water from the Skin and Lungs. In *The Journal of Biological Chemistry*. Vol. LXVII, No. 2. New York, 1925. s. 397-411.
- [33] ŘEZANKOVÁ, Hana; HÚSEK, Dušan; SNÁŠEL, Václav. *Shluková analýza dat*. Druhé rozšíření vydání. Praha : Professional Publishing, 2009. 218 s. ISBN 978-80-86946-81-8.
- [34] SAHA, I.; MAULIK, U.; NILANJAN; Differential fuzzy clustering for categorical data. In *Methods and Models in Computer Science, 2009. ICM2CS 2009. Proceeding of International Conference on* , vol., no., pp.1-6, 14-15 Dec. 2009
- [35] SHYUAN Wang; MINGQUAN Zhou; GUOHUA Geng; Application of fuzzy cluster analysis for medical image data mining. In *Mechatronics and Automation, 2005 IEEE International Conference* , vol.2, no., pp. 631- 636 Vol. 2, 29 July-1 Aug. 2005
- [36] SCHILLING, Christophe H.; LETSCHER, David; PALSSON, Bernhard. Theory for the Systemic Definition of Metabolic Pathways and their use in Interpreting Metabolic Function from a Pathway-Oriented Perspective. In *Journal of Theoretical Biology*. Volume 203, Issue 3. Academic Press, 2000. s. 229-248 . ISSN 0022-5193.

- [37] TAFTI, A.D.; SADATI, N.; Fuzzy clustering means data association algorithm using an adaptive neuro-fuzzy network. In *Aerospace conference*, 2009 IEEE , vol., no., pp.1-5, 7-14 March 2009
- [38] THAUNG THAUNG Win; LIN Mon; Document clustering by fuzzy c-mean algorithm. In *Advanced Computer Control (ICACC), 2010 2nd International Conference on* , vol.1, no., pp.239-242, 27-29 March 2010
- [39] THOMAS, B.; RAJU, G.; A fuzzy threshold based unsupervised clustering algorithm for natural data exploration. In *Networking and Information Technology (ICNIT), 2010 International Conference on* , vol., no., pp.473-477, 11-12 June 2010
- [40] TROJAN, Stanislav. *Lékařská fyziologie*. 2. přepracované a rozšířené vydání. Praha : Grada, 1996. 489 s. ISBN 978-80-7169-311-1.
- [41] TZUNG-PEI Hong; On Genetic-Fuzzy Data Mining Techniques. In *Granular Computing, 2007. GRC 2007. IEEE International Conference on* , vol., no., pp.3, 2-4 Nov. 2007
- [42] URBÁNKOVÁ, Eva. *Korelace poměru respirační výměny a hladiny glukózy během pohybové aktivity* [online]. Brno : Masarykova univerzita, 2010. 89 s. Diplomová práce. Masarykova univerzita, Fakulta sportovních studií. Dostupné z WWW: http://is.muni.cz/th/175824/fsps_m/Eva_Urbankova.pdf.
- [43] User's guide for QuadScan 4000 & MultiScan 5000. USA : BODYSTAT LTD., 2000. 86 s.
- [44] VANISRI, D.; Loganathan, C.; Fuzzy pattern cluster scheme for breast cancer datasets. In *Communication and Computational Intelligence (INCOCCI), 2010 International Conference on* , vol., no., pp.410-414, 27-29 Dec. 2010
- [45] WALENTA, Stefan, et al. Metabolic classification of human rectal adenocarcinomas: a novel guideline for clinical oncologists?. In *Journal of Cancer Research and Clinical Oncology*. 2003-06-01, 6, 129, s. 321-326. ISSN 0171-5216.
- [46] YANLI Zhang; YAN Ren; XIAODONG Liu; A New Method for Fuzzy Clustering Analysis Based on AFS Fuzzy Logic. In *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference on* , vol., no., pp.1-5, 12-14 Oct. 2008
- [47] ZHAOCHENG Hou; JIANMING Wang; A new cluster validity criterion for the cross iterative fuzzy clustering algorithm. In *Intelligent Control and Automation, 2004. WCICA 2004. Fifth World Congress on* , vol.3, no., pp. 2322- 2326 Vol.3, 15-19 June 2004

Seznam publikací a výsledků autora

Odborné články

1. PRAUZEK, Michal. Modern Methods of Measurement ECG. In *Lékař a technika - biomedicínské inženýrství a informatika : YBERC 2008*. Praha : Česká lékařská společnost J. E. Purkyně a ČVUT Fakulta biomedicínského inženýrství, 2008. s. 215. ISSN 0301-5419. (100%)
2. ADAMEC, Ondřej, PRAUZEK, Michal. Snímání srdeční aktivity v mobilních systémech. In *Sborník přednášek III. Bilaterální symposium v Rožnově pod Radhoštěm, 16.-18.4.2008, Rožnov pod Radhoštěm, ČR*. Ostrava : Česká společnost pro úrazovou chirurgii ČLS JEP, 2008. ISBN 978-80-254-1663-1. (50%)
3. PRAUZEK, Michal; PENHAKER, Marek; BERNABUCCI, Ivan; CONFORTO, Silvia. ECG - Precordial Leads Reconstruction. In *9th International Conference on Information Technology and Applications in Biomedicine : Final Program and Abstract Book*. Larnaca, Cyprus : IEEE, 2009. s. 71. ISBN 978-1-4244-5379-5. (85%)
4. PRAUZEK, Michal; PENHAKER, Marek. Methods of Comparing ECG Reconstruction. In *2009 2nd International Conference on Biomedical Engineering and Informatic*. Tianjin, China : IEEE, 2009. s. 675-678. ISBN 978-1-4244-4133-4. (95%)
5. PRAUZEK, Michal; CONFORTO, Silvia. Component Analysis of ECG. In *Trendy v biomedicínském inženýrství. 16 18. Září 2009*. Bratislava : Slovenská technická univerzita v Bratislave, 2009. ISBN 978-80-227-3105-8 (90%)
6. PRAUZEK, Michal. Modern Measurement of Precordial ECG Leads. In *Ph.D. Workshop of Faculty of Electrical Engineering and Computer Science (WOFEX09), 8.- 9. 9.2009*, Ostrava : VŠB - Technical University of Ostrava, FEECS, 2009. ISBN: 978-80-248-2028-6 (100%)
7. PRAUZEK, Michal; ADAMEC, Ondřej; PENHAKER, Marek. Testing of Portable Telemedical Power Supply. In *Proceeding of 2010 International Conference on Signal Processing and System : Volume 2*. Dalian, China : IACSIT, 2010. s. 380-382. ISBN 978-1-4244-6891-1. (90%)
8. PRAUZEK, Michal; PETEREK, Tomáš; ADAMEC, Ondřej; STANKUŠ, Martin; PENHAKER, Marek. Simple Data Aquisition System for Photoplethysmography and Electrocardiography. In *Proceeding of 2010 International Conference on Signal Processing and System : Volume 2*. Dalian, China : IACSIT, 2010. s. 377-379. ISBN 978-1-4244-6891-1. (50%)

9. PETEREK Tomáš; PRAUZEK, Michal; PENHAKER, Marek. Baseline Wander Elimination by Fourier Series. In *Proceeding of 2010 International Conference on Signal Processing and System : Volume 1*. Dalian, China : IACSIT, 2010. s. 365-367. ISBN 978-1-4244-6891-1. (40%)
10. PETEREK Tomáš; PRAUZEK, Michal; PENHAKER, Marek. A New Method for Identification of the Significant Point in the Plethysmographical Record. In *Proceeding of 2010 International Conference on Signal Processing and System : Volume 1*. Dalian, China : IACSIT, 2010. s. 362-364. ISBN 978-1-4244-6891-1. (40%)
11. PRAUZEK, Michal; PETEREK Tomáš; ADAMEC, Ondřej; PENHAKER, Marek. Analog Signal Preprocessing in Reflected Plethysmography. In *Proceeding of 2010 International Conference on Signal Processing and System : Volume 1*. Dalian, China : IACSIT, 2010. s. 358-361. ISBN 978-1-4244-6891-1. (50%)
12. PRAUZEK, Michal; PETEREK Tomáš; PENHAKER, Marek. Reduction of ECG Pre-cordial Leads. In *Proceeding of 2010 International Conference on Signal Processing and System : Volume 1*. Dalian, China : IACSIT, 2010. s. 354-357. ISBN 978-1-4244-6891-1. (50%)
13. PRAUZEK, Michal; PETEREK, Tomáš; ADAMEC, Ondřej; PENHAKER, Marek. Device of Accurate Reflex Plethysmograph. In *7th ESBME: Proceedings European Symposium on Biomedical Engineering*. Chalkidiki, Greece, 2010. s. 111. ISBN 978-960-85715-5-6. (90%)
14. ČERNÝ, Martin; ŽŮREK, Petr; PRAUZEK, Michal; KREJCAR, Ondřej; PENHAKER, Marek. New approaches for continuous non invasive blood pressure monitoring. In *12th Mediterranean Conference on Medical and Biological Engineering and Computing*. 1. Chalkidiki : IEEE, 2010. s. 228-231. (20%)

Předměty průmyslového vlastnictví

1. PRAUZEK, Michal; PENHAKER, Marek; ADAMEC, Ondřej. Užitený vzor: *Reflexní fotopletysmograf*. VŠB – TU Ostrava, 2010. Ev.č. 005/22-2-2010, číslo přihlášky PUV-2010-23335. (80%)
2. PRAUZEK, Michal; PENHAKER, Marek; ADAMEC, Ondřej. Užitený vzor: *Napájecí zdroj pro jednočlávkové aplikace*. VŠB – TU Ostrava, 2010. Ev.č. 006/22-2-2010, číslo přihlášky PUV-2010-23334. (80%)

Evidence vědy a výzkumu

1. PRAUZEK, Michal; PENHAKER, Marek; ADAMEC, Ondřej. Funkční vzorek: *Reflexní fotopletysmograf*. VŠB – TU Ostrava, 2010. Ev.č. 001/22-2-2010.F. (80%)
2. PRAUZEK, Michal; PENHAKER, Marek; ADAMEC, Ondřej. Funkční vzorek: *Napájecí zdroj pro jednočlávkové aplikace*. VŠB – TU Ostrava, 2010. Ev.č. 002/22-2-2010.F. (80%)
3. PRAUZEK, Michal; PETEREK, Tomáš; PENHAKER, Marek. Software: *ECG DRIFT*. VŠB – TU Ostrava, 2010. Ev.č. 002/22-2-2010_SW. (40%)

4. PRAUZEK, Michal; MICHALÍKOVÁ Markéta; PENHAKER, Marek. Funkční vzorek: *Modul pro evaluaci EKG*. VŠB – TU Ostrava, 2011. Ev.č. 036/27-05-2011_F. (50%)
5. PRAUZEK, Michal; HLAVICA Jakub; PENHAKER, Marek. Software: *Nástroj pro detekci metabolických vazeb*. VŠB – TU Ostrava, 2011. Ev.č. 006/26-05-2011_SW. (50%)
6. PRAUZEK, Michal; RAJMANOVÁ Petra; PENHAKER, Marek. Software: *Systém pro automatické hodnocení glykémie*. VŠB – TU Ostrava, 2011. Ev.č. 00/26-05-2011_SW. (50%)

Řešené granty

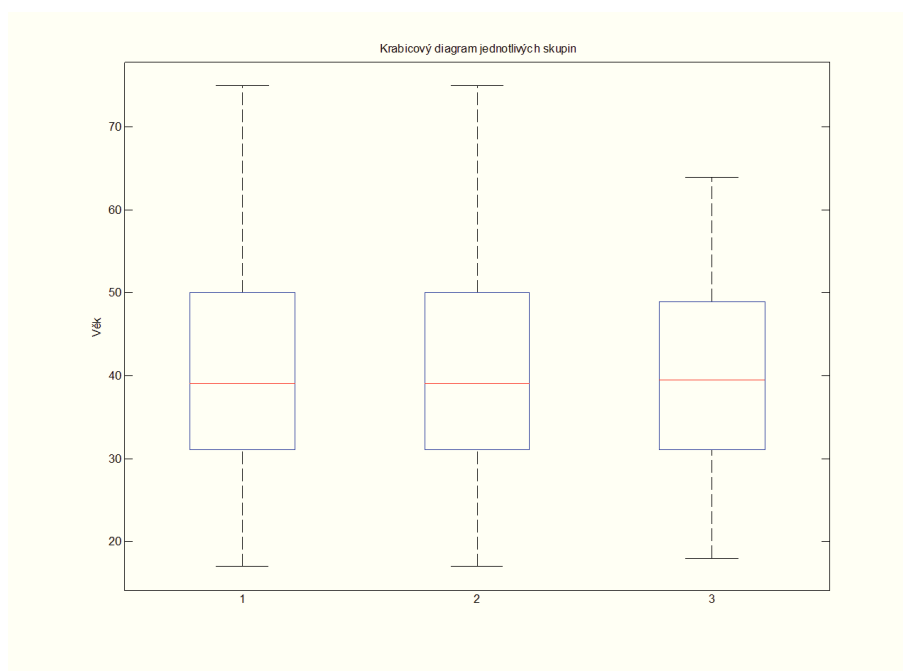
1. ČERNOHORSKÝ, Jindřich; PENHAKER, Marek. *Biomedicínské inženýrské systémy III*. VŠB – TU Ostrava, 2007. Interní grantový projekt.
2. ČERNOHORSKÝ, Jindřich; PENHAKER, Marek. *Biomedicínské inženýrské systémy IV*. VŠB – TU Ostrava, 2008. Interní grantový projekt.
3. PRAUZEK, Michal; SROVNAL, Vilém; PENHAKER, Marek; HORÁK, David. *Parametrické modely šíření EKG na povrchu hrudníku*. VŠB – TU Ostrava, 2009. IGA 14/2009.
4. PENHAKER, Marek: SGS Studentská grantová soutěž 2010, *Biomedicínské inženýrské systémy VI*. SP/2010190, int. Číslo 14512. VŠB – TU Ostrava, 2010. Interní grantový projekt.
5. PENHAKER, Marek; ČERNÝ, Martin; AUGUSTYNEK, Martin: *Biomedicínská technika na středních školách – CZ.1.07/1.1.07/02.0075*, Zvyšování kvality ve vzdělávání v kraji Moravskoslezském Operačního programu: Vzdělávání pro konkurenceschopnost.
6. SROVNAL, Vilém: *Bezpečnost a zabezpečení aplikací sítí vestavěných systémů*. GAČR 102/08/1429
7. ČERNOHORSKÝ, Jindřich: CAK 1M0567 CAK Centrum Aplikované Kybernetiky
8. ČERNÝ Martin, PENHAKER Marek, PRAUZEK Michal: *Inovace předmětu Biotelemetrie*. FRVŠ 1504/2011, kategorie F3

Získaná ocenění

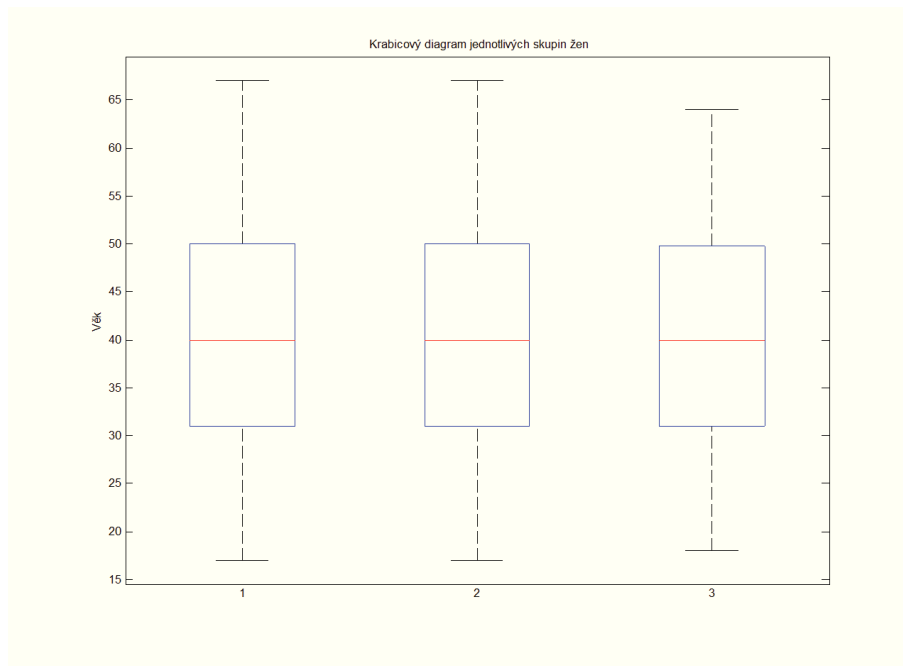
1. STOČ 2008, Univerzita Tomáše Bati ve Zlíně, 2. místo kategorie S4 – Aplikace měřicích a diagnostických systémů
2. Trendy v Biomedicínském Inženýrství 2009, STU Bratislava, 1. místo za nejlepší prezentaci konference, cena „Young Investigator Award“.
3. Udělení stipendia Statutárního města Ostravy za dosažení předepsaného studijního průměru, zapojení do vědecké činnosti, prezentace města a veřejně prospěšné činnosti v akademickém roce 2010/2011.

Příloha A

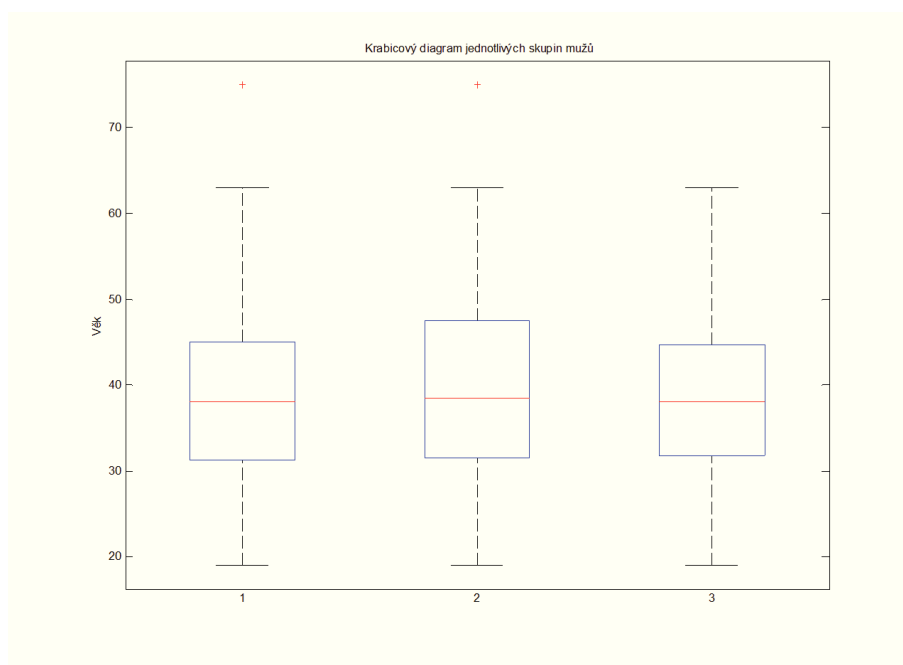
Rozdělení skupin pro analýzu a testování



Obrázek A.1: Krabicový diagram: celkem, analýza a testování všech



Obrázek A.2: Krabicový diagram: celkem, analýza a testování žen



Obrázek A.3: Krabicový diagram: celkem, analýza a testování mužů

Příloha B

Analýza hlavních komponent

Tabulka B.1: Redukce testu A

Parametr	Součet	V redukci	Zdůvodnění odmítnutí/přijetí
BMI	0,6956	ANO	Zadáno lékařem
RQ1	0,7732	NE	Obsaženo ve fat1
RQ2	1,0438	NE	Obsaženo ve fat2
RQ3	0,7643	NE	Obsaženo ve fat3
fat1	0,8574	ANO	Zadáno lékařem
fat2	1,0183	ANO	Zadáno lékařem
fat3	0,7903	ANO	Zadáno lékařem
carbohydrate1	0,7045	NE	Informace < 0,8
carbohydrate2	0,3726	NE	Informace < 0,8
carbohydrate3	0,2439	NE	Informace < 0,8
insulin1	0,9585	ANO	-
c-peptid1	1,0656	ANO	-
insulin2	1,0317	ANO	-
c-peptid2	0,9828	ANO	-
insulin3	1,1151	ANO	-
c-peptid3	1,0981	ANO	-
alt	0,7479	NE	Informace < 0,8
ast	0,4095	NE	Informace < 0,8
gmt	0,5644	NE	Informace < 0,8
moč	0,7724	NE	Informace < 0,8
tag	0,5566	NE	Informace < 0,8
ldlv	0,5221	NE	Informace < 0,8
hdlc	0,6116	NE	Informace < 0,8
chol	0,3283	NE	Informace < 0,8
glu0	0,6913	NE	Informace < 0,8
glu1	0,7019	NE	Informace < 0,8
glu2	0,6620	NE	Informace < 0,8

Tabulka B.2: Redukce testu B

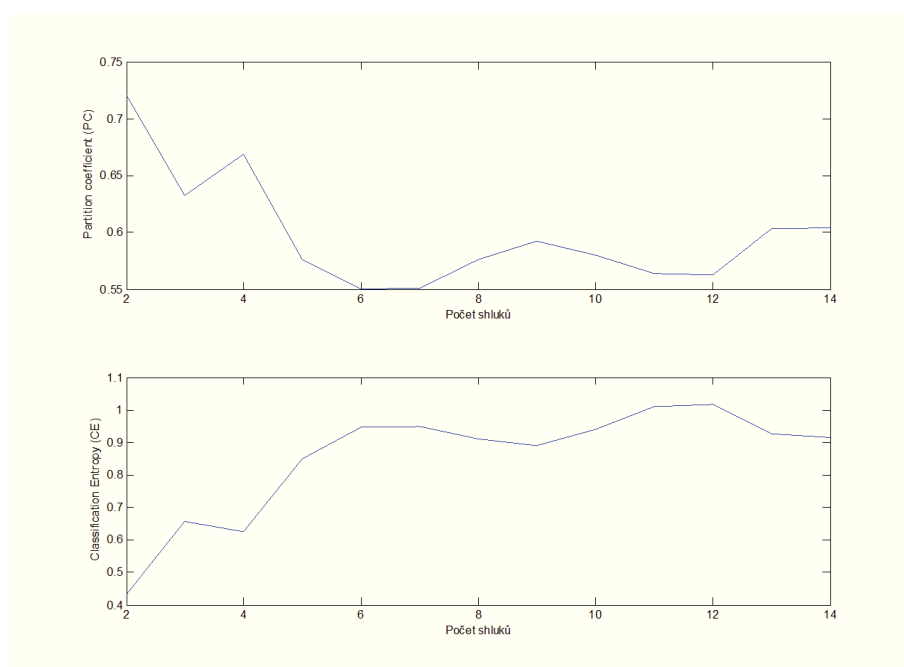
Parametr	Součet	V redukci	Zdůvodnění odmítnutí/přijetí
BMI	1,1231	ANO	-
insulin1	0,9996	ANO	-
c-peptid1	1,1087	ANO	-
insulin2	0,9364	ANO	-
c-peptid2	0,8434	ANO	-
insulin3	1,1062	ANO	-
c-peptid3	1,1062	ANO	-
tuk	1,3730	ANO	-
ath	1,3727	ANO	-
voda	1,3589	ANO	-
bezvody	0,2151	NE	Informace < 0,8
ict	0,8368	ANO	-
ect	0,8783	ANO	-
cvt	1,3214	ANO	-
alt	0,8682	ANO	-
ast	0,5697	NE	Informace < 0,8
gmt	0,7510	NE	Informace < 0,8
moč	0,7926	NE	Informace < 0,8
tag	0,6486	NE	Informace < 0,8
ldlv	0,5527	NE	Informace < 0,8
hdlc	0,4891	NE	Informace < 0,8
chol	0,4560	NE	Informace < 0,8
glu0	0,6594	NE	Informace < 0,8
glu1	0,8861	ANO	-
glu2	0,9114	ANO	-

Tabulka B.3: Redukce testu C

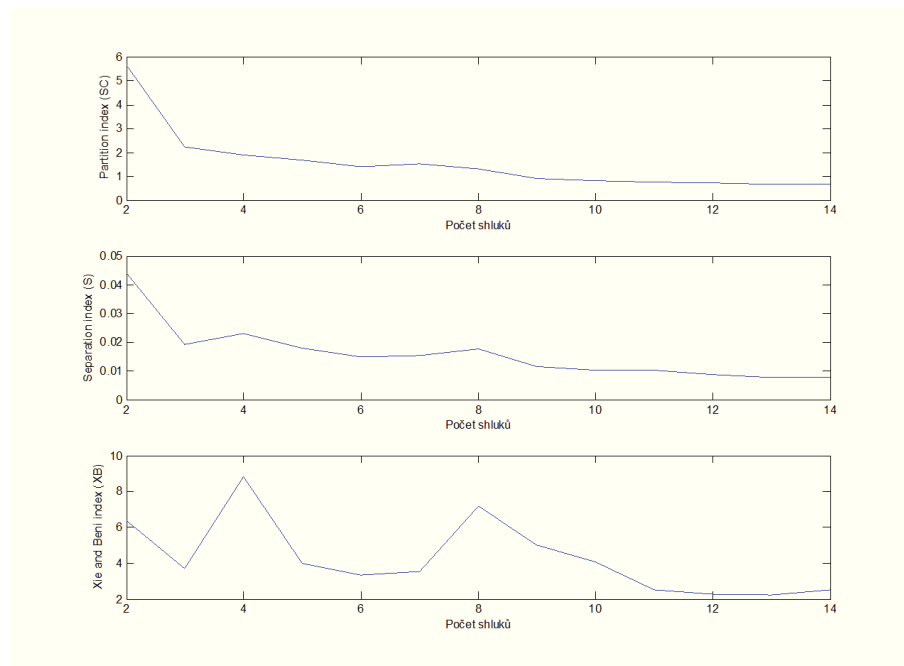
Parametr	Součet	V redukci	Zdůvodnění odmítnutí
BMI	1,5776	ANO	-
RQ1	1,3127	NE	Obsaženo ve fat1
RQ2	1,3476	NE	Obsaženo ve fat2
RQ3	1,1898	NE	Obsaženo ve fat3
fat1	1,3515	ANO	-
fat2	1,4217	ANO	-
fat3	1,3239	ANO	-
carbohydrate1	1,2420	NE	Redundantní – počítáno z fat1
carbohydrate2	0,5226	NE	Redundantní – počítáno z fat2
carbohydrate3	0,4447	NE	Redundantní – počítáno z fat3
insulin1	1,3274	ANO	-
c-peptid1	1,3835	ANO	-
insulin2	1,2421	ANO	-
c-peptid2	1,0994	ANO	-
insulin3	1,4720	ANO	-
c-peptid3	1,5888	ANO	-
tuk	1,5561	ANO	-
ath	1,5543	ANO	-
voda	1,5935	ANO	-
bezvody	1,2546	ANO	-
ict	0,9214	NE	Informace < 1
ect	1,1802	ANO	-
cvt	1,5851	ANO	-
alt	1,2037	ANO	-
ast	0,6441	NE	Informace < 1
gmt	0,8583	NE	Informace < 1
moč	1,3810	ANO	-
tag	0,7988	NE	Informace < 1
ldlv	1,0833	ANO	-
hdlc	1,2218	ANO	-
chol	1,0917	ANO	-
glu0	0,9377	NE	Informace < 1
glu1	1,1958	ANO	-
glu2	1,3641	ANO	-

Příloha C

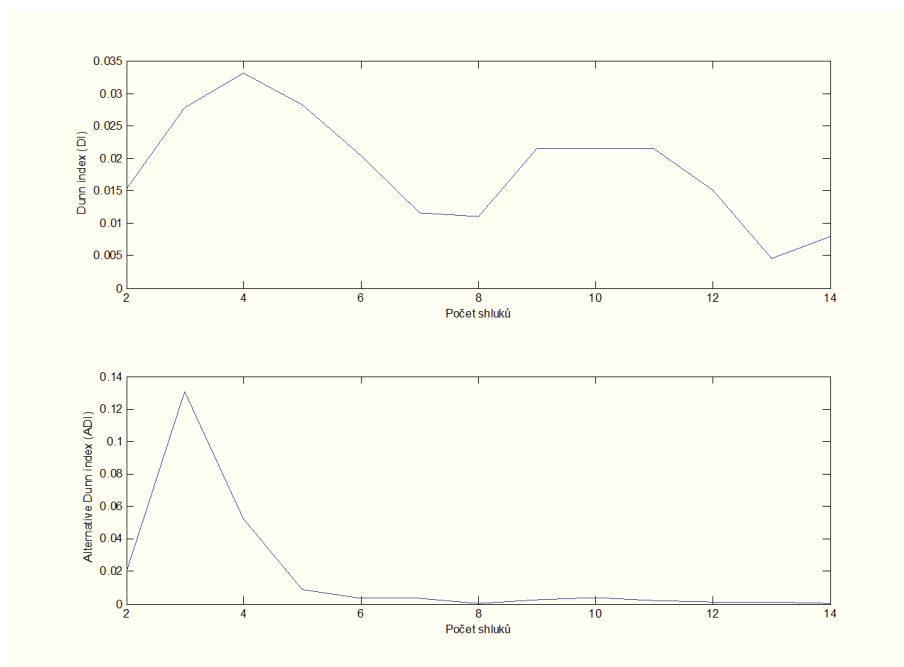
Stanovení počtu shluků



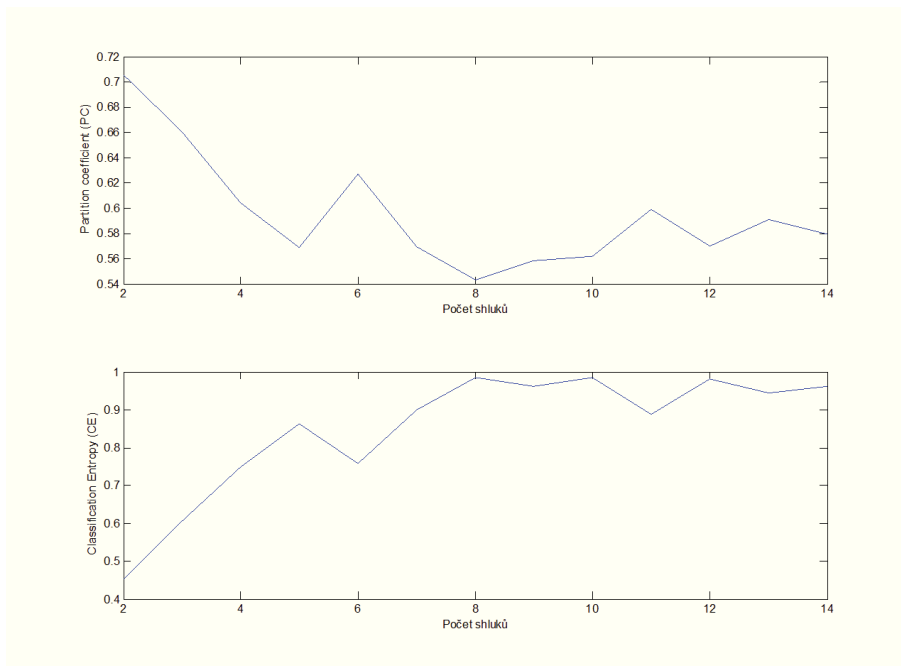
Obrázek C.1: Indexy počtu shluků PC a CE - test A



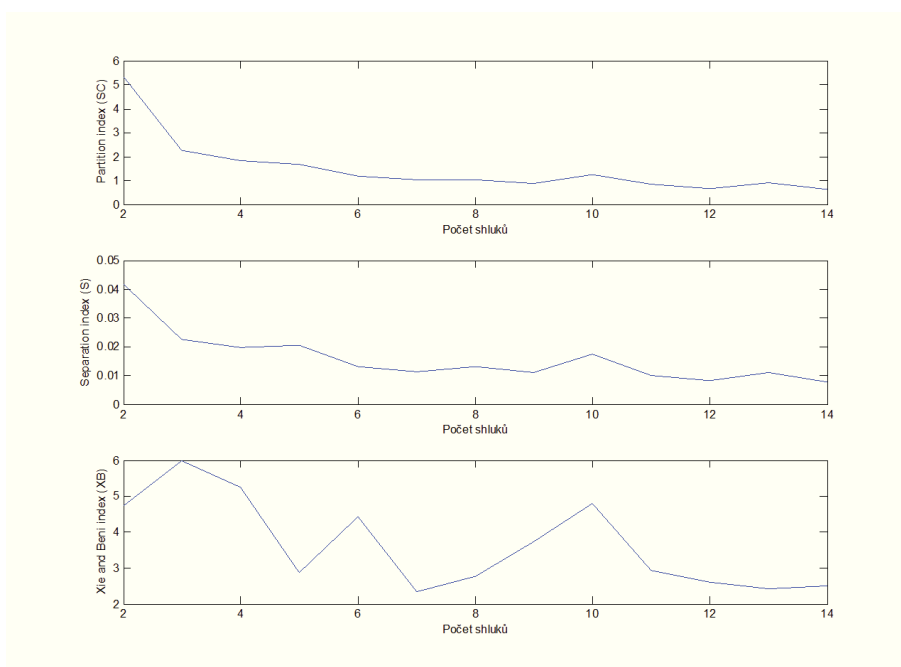
Obrázek C.2: Indexy počtu shluků SC, S a XB - test A



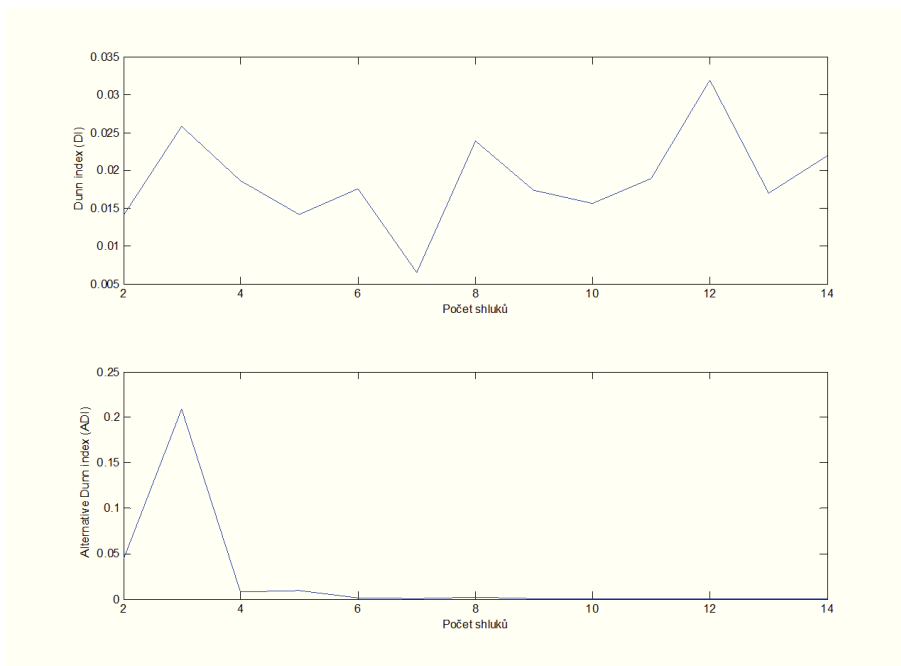
Obrázek C.3: Indexy počtu shluků DI a ADI - test A



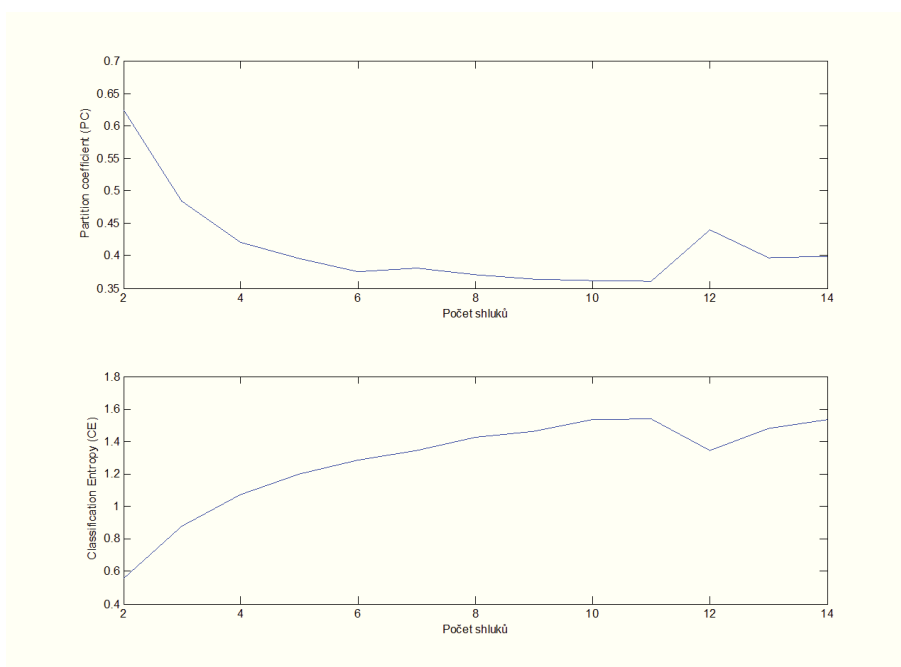
Obrázek C.4: Indexy počtu shluků PC a CE - test B



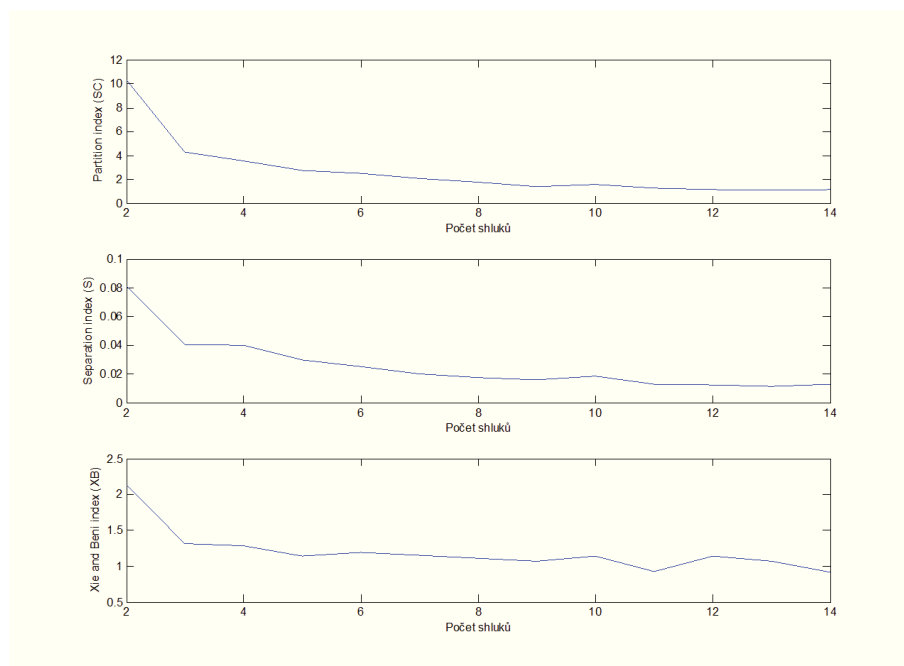
Obrázek C.5: Indexy počtu shluků SC, S a XB - test B



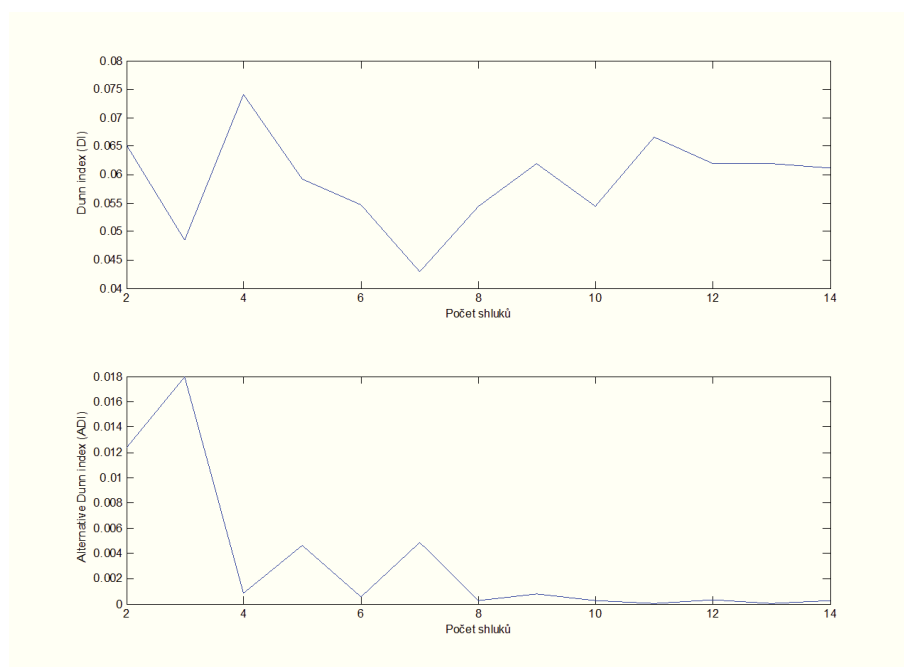
Obrázek C.6: Indexy počtu shluků DI a ADI - test B



Obrázek C.7: Indexy počtu shluků PC a CE - test C



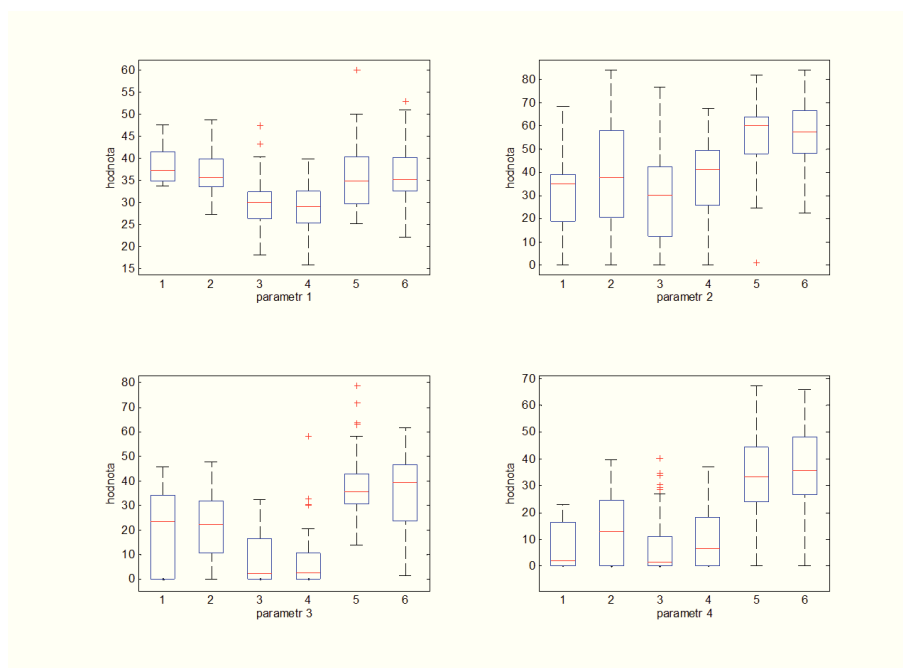
Obrázek C.8: Indexy počtu shluků SC, S a XB - test C



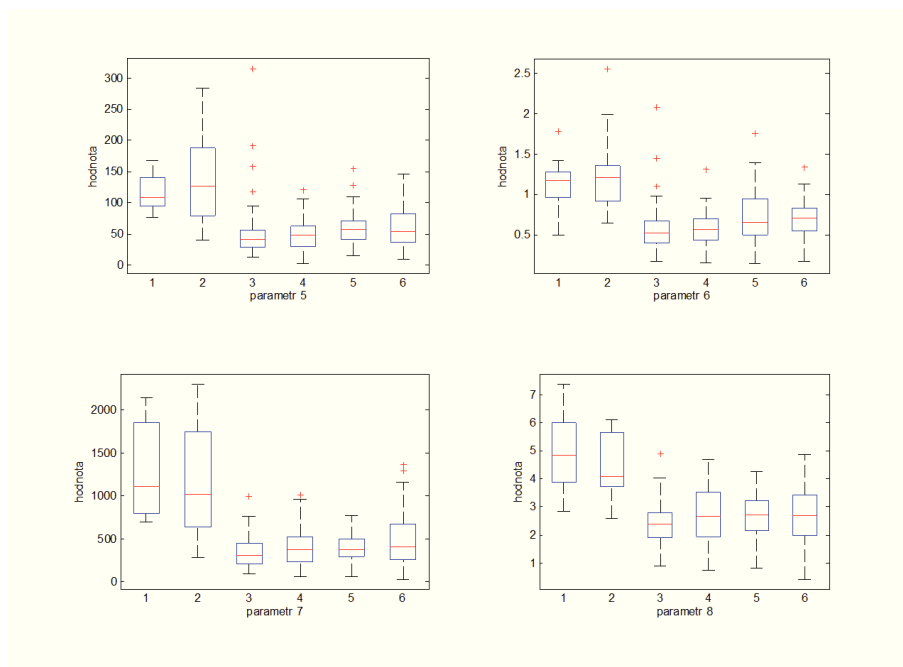
Obrázek C.9: Indexy počtu shluků DI a ADI - test C

Příloha D

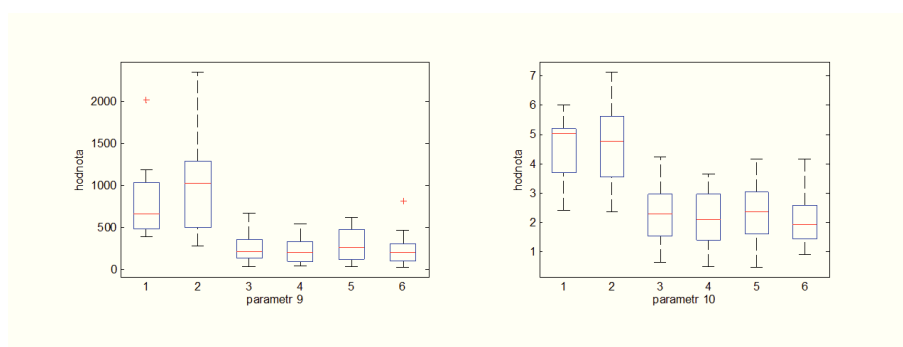
Grafická evaluace algoritmu



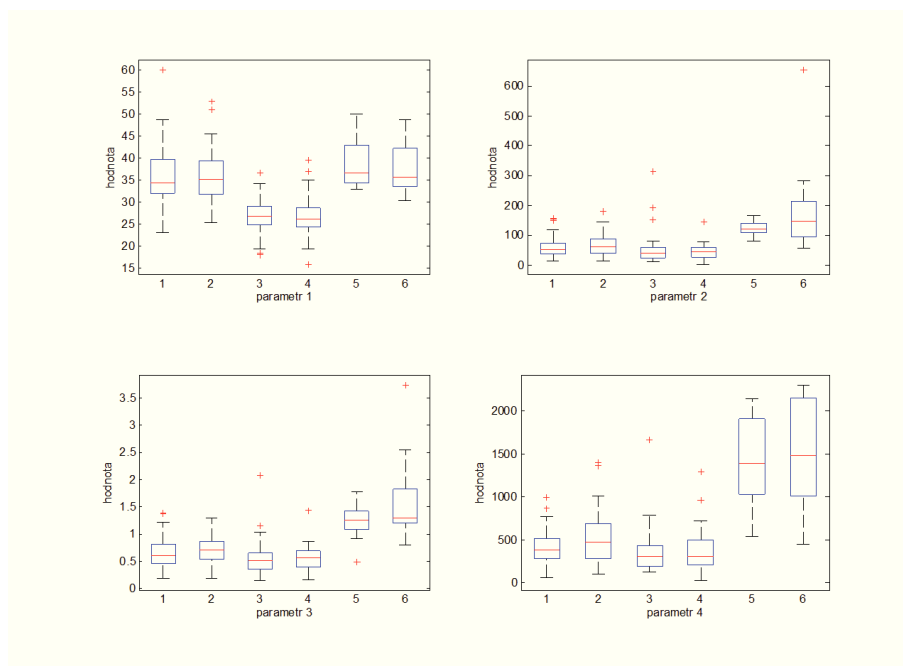
Obrázek D.1: První 4 parametry klasifikace testu A



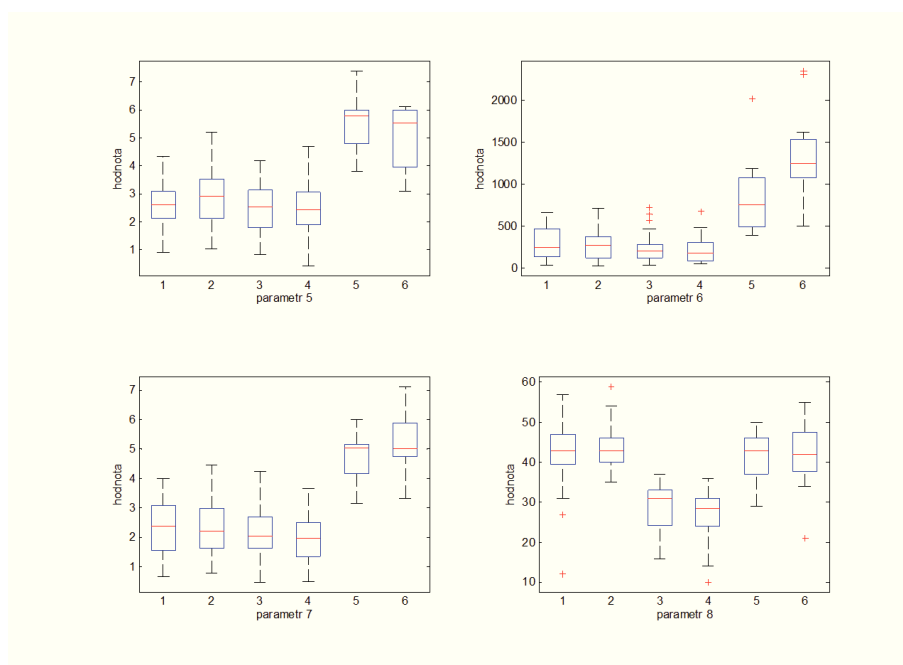
Obrázek D.2: Druhé 4 parametry klasifikace testu A



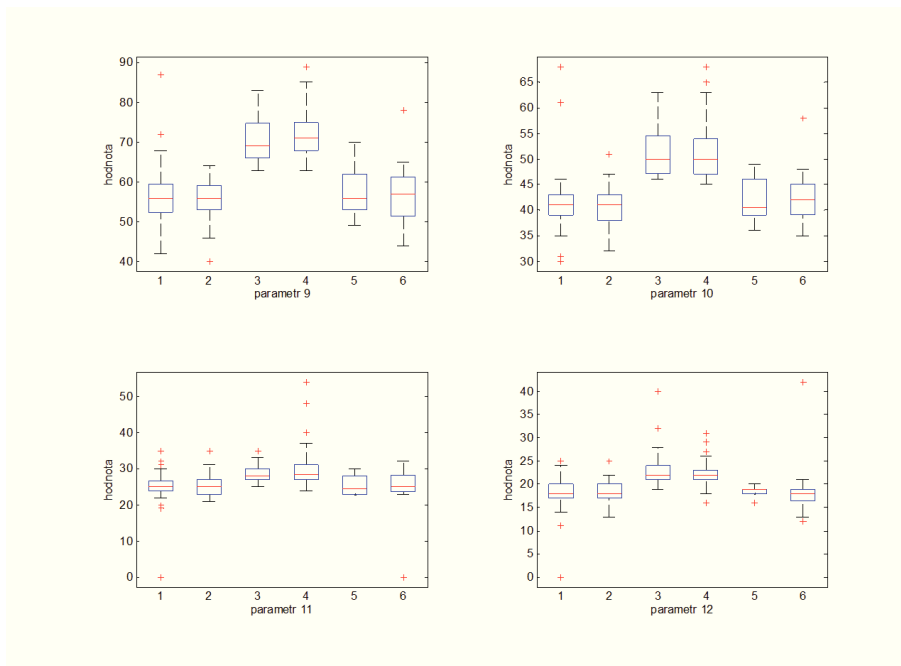
Obrázek D.3: Poslední 2 parametry klasifikace testu A



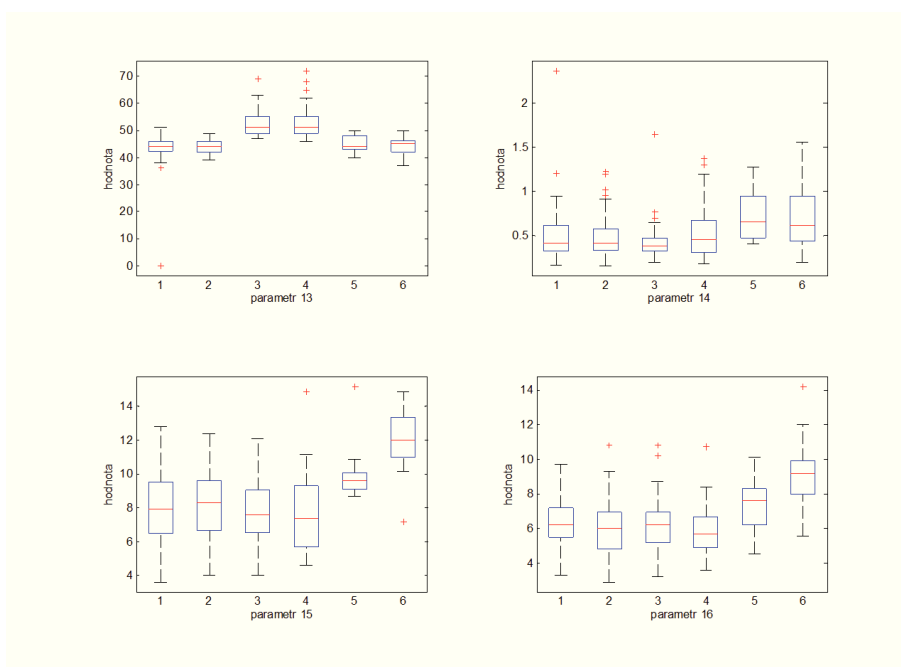
Obrázek D.4: První 4 parametry klasifikace testu B



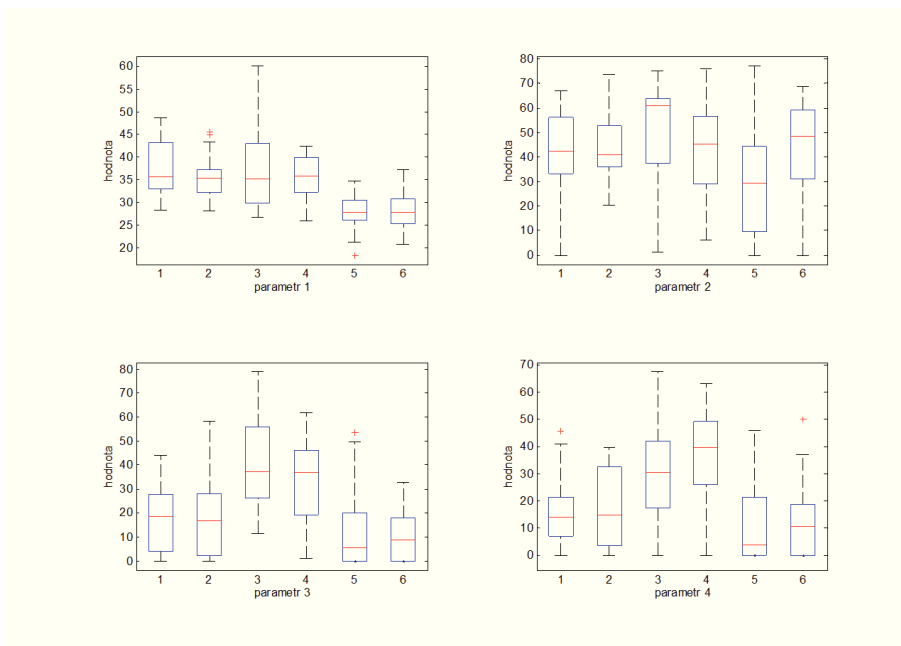
Obrázek D.5: Druhé 4 parametry klasifikace testu B



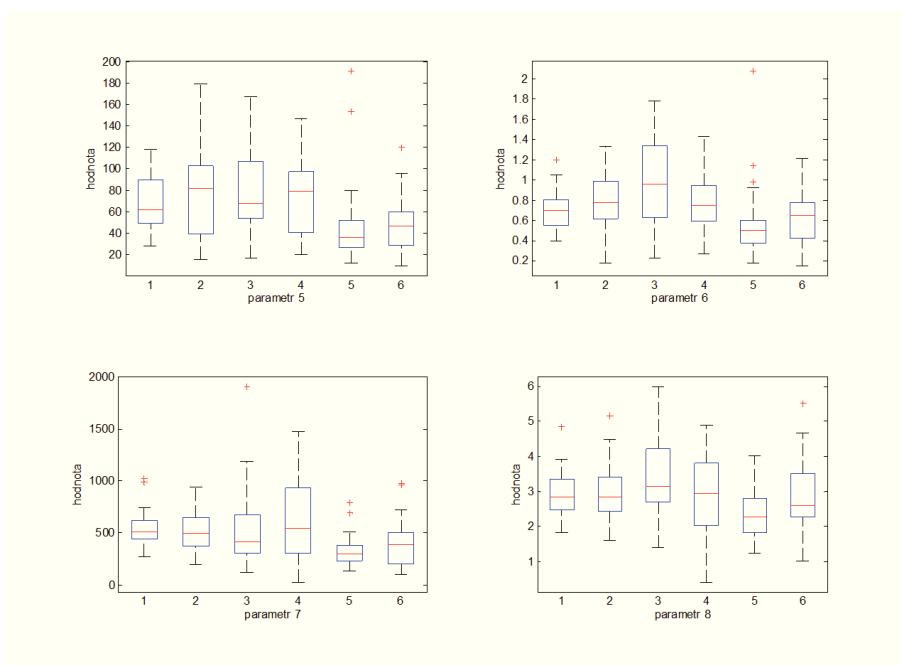
Obrázek D.6: Třetí 4 parametry klasifikace testu B



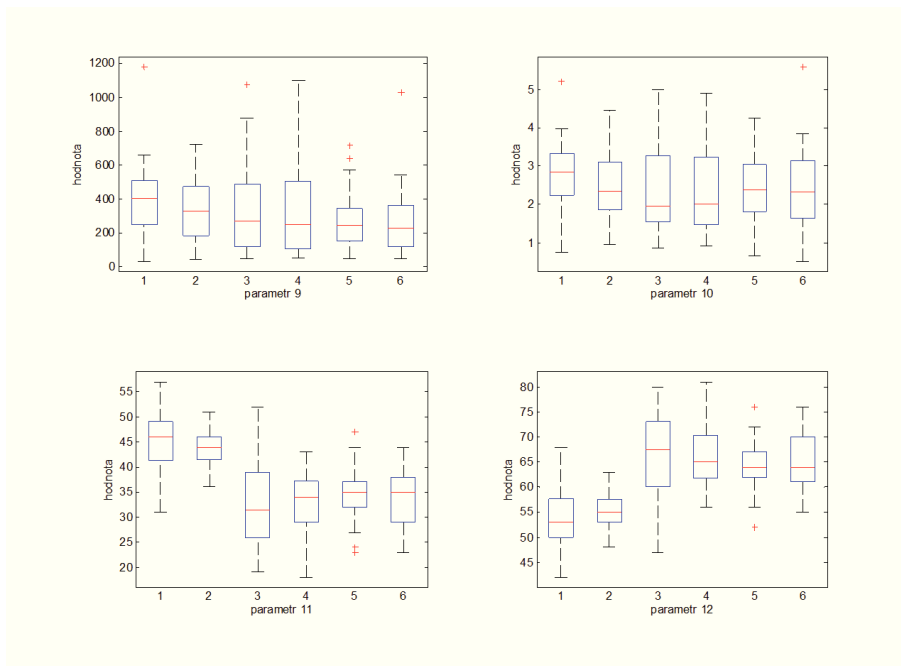
Obrázek D.7: Poslední 4 parametry klasifikace testu B



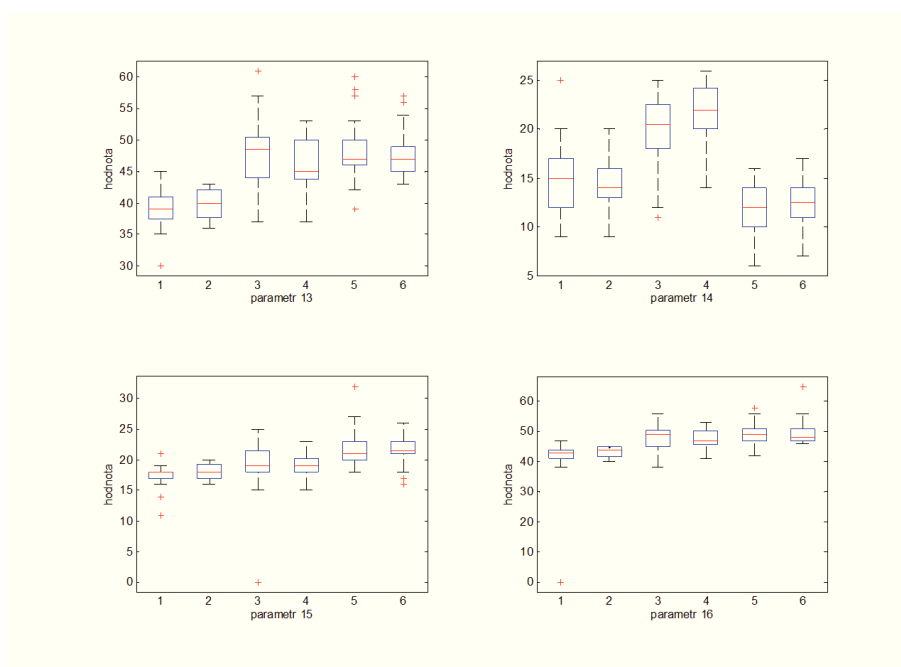
Obrázek D.8: První 4 parametry klasifikace testu C



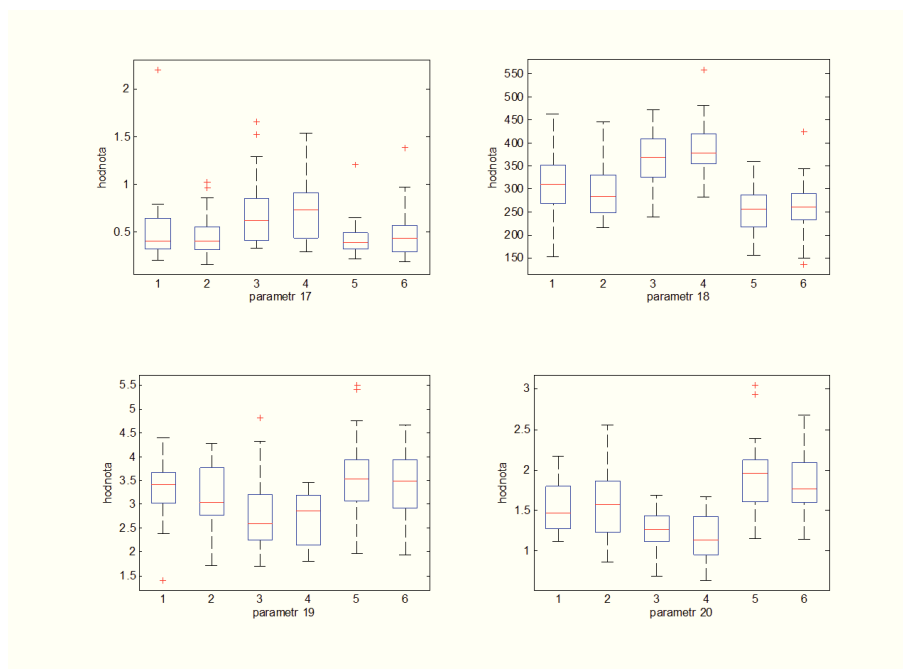
Obrázek D.9: Druhé 4 parametry klasifikace testu C



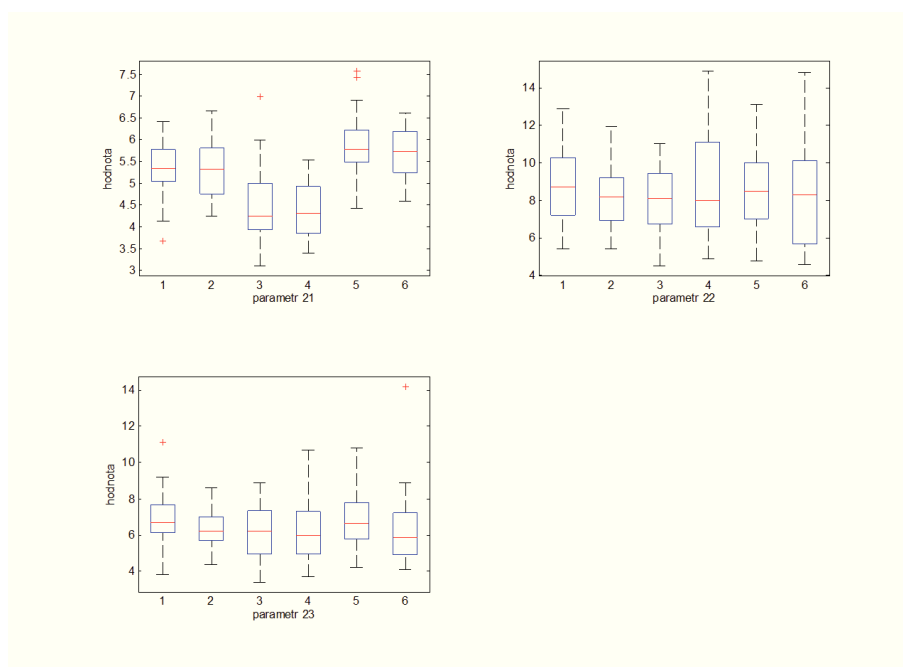
Obrázek D.10: Třetí 4 parametry klasifikace testu C



Obrázek D.11: Čtvrté 4 parametry klasifikace testu C



Obrázek D.12: Páté 4 parametry klasifikace testu C



Obrázek D.13: Poslední 3 parametry klasifikace testu C

Příloha E

Statistická evaluace výsledků

V této příloze se nacházejí tabulky, které popisují statistické testování hypotéz a jednotlivé statistické parametry testovaných veličin. U testování hypotéz je vždy uveden výsledek testování, kdy nula znamená, že nulová hypotéza byla přijata. Zkratka A-F, B-F a C-F značí výsledky souhlasného testování shluku A, B a C pomocí F-testu. Ke každému testování je přiřazena p hodnota označená A-p, B-p a C-p, která se vztahuje k řádku o jednu pozici výše. Obdobně jsou uvedeny hodnoty pro T-test.

Jsou zde rovněž uvedeny výsledky křížového testování. Systém značení je obdobný jako u souhlasného testování, označení AB-F, AC-F a BC-F značí testování hypotézy mezi jednotlivými shluky AB, AC a BC. Ke každému testování je přiřazena p hodnota označená jako AB-p, AC-p a BC-p, která se vztahuje k předchozímu řádku. Pro T-test je systém stejný, výsledky jsou označeny jako AB-T, AC-T a BC-T pro dvojice shluků AB, AC a BC.

Ve zbývajících tabulkách jsou uvedeny podobné statistické parametry jednotlivých veličin ve shlucích. Označení A1, B1 a C2 náleží analytickým shlukům A, B a C, testovací shluky jsou označeny jako A2, B2 a C2. V těchto tabulkách se vyskytují 3 základní statistické hodnoty a to medián, průměrná hodnota a střední kvadratická odchylka.

Tabulka E.1: Testování hypotéz – souhlasný test A

	BMI	fat1	fat2	fat3	ins1	c-p1	ins2	c-p2	ins3	c-p3
A-F	0	0	0	0	1	0	0	0	0	0
A-p	0,57	0,46	0,33	0,37	0,02	0,28	0,57	0,44	0,59	0,77
A-T	0	0	0	0	0	0	0	0	0	0
A-p	0,42	0,24	0,79	0,11	0,27	0,53	0,57	0,37	0,36	0,88
B-F	0	0	0	0	1	1	1	1	0	0
B-p	0,90	0,19	0,64	0,93	0,00	0,05	0,02	0,04	0,28	0,37
B-T	0	0	0	0	0	0	0	0	0	0
B-p	0,49	0,06	0,45	0,20	0,70	0,79	0,16	0,06	0,63	0,50
C-F	0	0	0	0	0	1	1	0	0	0
C-p	0,21	0,26	0,78	0,27	0,79	0,02	0,00	0,34	0,26	0,28
C-T	0	0	0	0	0	0	0	0	0	0
C-p	0,97	0,67	0,48	0,45	0,95	0,67	0,07	0,74	0,10	0,27

Tabulka E.2: Testování hypotéz – křížový test A

	BMI	fat1	fat2	fat3	ins1	c-p1	ins2	c-p2	ins3	c-p3
AB-F	0	0	1	0	0	0	1	1	1	0
AB-p	0,63	0,65	0,01	0,82	0,25	0,64	0,00	0,00	0,00	0,38
AB-T	1	0	0	0	1	1	1	1	1	1
AB-p	0,00	0,77	0,06	0,99	0,00	0,00	0,00	0,00	0,00	0,00
AC-F	0	0	0	1	0	0	1	1	1	0
AC-p	0,07	0,56	0,55	0,04	0,86	0,68	0,00	0,04	0,00	0,41
AC-T	0	1	1	1	1	1	1	1	1	1
AC-p	0,26	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00
BC-F	1	0	1	1	1	0	0	0	0	0
BC-p	0,01	0,10	0,01	0,00	0,03	0,14	0,85	0,51	0,50	0,99
BC-T	1	1	1	1	0	1	0	0	0	0
BC-p	0,00	0,00	0,00	0,00	0,27	0,03	0,33	0,16	0,39	0,95

Tabulka E.3: Statistické parametry shluků – test A

	BMI	fat1	fat2	fat3	ins1	c-p1	ins2	c-p2	ins3	c-p3
$\tilde{x}(\mathbf{A1})$	37,2	35,0	23,4	2,0	108	1,17	1104	4,84	659	5,05
$\tilde{x}(\mathbf{B1})$	30,1	30,2	2,3	1,5	41	0,52	312	2,38	208	2,28
$\tilde{x}(\mathbf{C1})$	34,9	60,1	35,8	33,5	57	0,65	367	2,72	262	2,35
$\tilde{x}(\mathbf{A2})$	35,6	37,9	22,4	12,6	127	1,21	1018	4,10	1028	4,78
$\tilde{x}(\mathbf{B2})$	29,0	41,2	2,5	6,5	49	0,57	372	2,68	197	2,12
$\tilde{x}(\mathbf{C2})$	35,2	57,6	39,2	35,8	53	0,71	401	2,69	197	1,93
$\bar{x}(\mathbf{A1})$	38,4	31,8	20,3	7,4	115	1,15	1311	4,95	833	4,51
$\bar{x}(\mathbf{B1})$	29,6	30,0	8,7	7,4	51	0,56	348	2,44	249	2,28
$\bar{x}(\mathbf{C1})$	36,2	55,6	38,8	33,5	60	0,73	383	2,69	283	2,27
$\bar{x}(\mathbf{A2})$	37,0	40,3	21,9	13,7	133	1,23	1189	4,52	1003	4,58
$\bar{x}(\mathbf{B2})$	29,0	36,6	7,3	9,9	49	0,57	402	2,75	235	2,18
$\bar{x}(\mathbf{C2})$	36,1	57,1	36,3	36,6	59	0,69	493	2,76	214	2,04
$\sigma^2(\mathbf{A1})$	21,7	324,4	302,4	92,3	1015	0,10	301690	1,87	224799	1,22
$\sigma^2(\mathbf{B1})$	29,4	432,5	110,4	109,9	1987	0,09	29284	0,59	28903	0,85
$\sigma^2(\mathbf{C1})$	65,1	252,9	233,2	324,7	971	0,14	27347	0,72	35054	0,84
$\sigma^2(\mathbf{A2})$	30,9	508,9	187,8	159,6	4489	0,20	432597	1,30	316309	1,48
$\sigma^2(\mathbf{B2})$	28,4	305,4	124,4	112,1	616	0,05	54630	1,00	21749	0,67
$\sigma^2(\mathbf{C2})$	42,7	171,9	212,3	223,5	890	0,06	104918	1,00	23858	0,58

Tabulka E.4: Testování hypotéz – souhlasný test B, část 1

	BMI	ins1	c-p1	ins2	c-p2	ins3	c-p3	tuk
A-F	0	0	0	1	1	0	0	1
A-p	0,07	0,84	0,20	0,00	0,02	0,23	0,79	0,02
A-T	0	0	0	1	1	0	0	0
A-p	0,88	0,41	0,29	0,02	0,01	0,37	0,72	0,55
B-F	0	1	0	0	0	0	0	0
B-p	0,17	0,00	0,11	0,95	0,15	0,66	0,45	0,80
B-T	0	0	0	0	0	0	0	0
B-p	0,96	0,48	0,99	0,80	0,86	0,69	0,53	0,22
C-F	0	1	0	0	0	0	0	0
C-p	0,90	0,00	0,06	0,89	0,87	0,99	0,64	0,66
C-T	0	0	0	0	0	1	0	0
C-p	0,81	0,13	0,20	0,69	0,27	0,04	0,16	0,78

Tabulka E.5: Testování hypotéz – souhlasný test B, část 2

	ath	voda	ict	ect	ctv	alt	glu1	glu2
A-F	1	1	1	1	1	1	0	0
A-p	0,02	0,00	0,00	0,00	0,00	0,05	0,92	0,90
A-T	0	0	0	0	0	0	0	0
A-p	0,59	0,53	0,36	0,44	0,41	0,87	0,66	0,16
B-F	0	0	1	0	0	0	0	0
B-p	0,82	0,38	0,00	0,13	0,10	0,06	0,35	0,76
B-T	0	0	0	0	0	0	0	0
B-p	0,24	0,87	0,17	0,26	0,65	0,06	0,79	0,49
C-F	0	0	1	1	0	0	0	0
C-p	0,66	0,50	0,01	0,00	0,89	0,24	0,90	0,43
C-T	0	0	0	0	0	0	1	1
C-p	0,78	0,86	0,78	0,87	0,64	0,90	0,02	0,01

Tabulka E.6: Testování hypotéz – křížový test B, část 1

	BMI	ins1	c-p1	ins2	c-p2	ins3	c-p3	tuk
AB-F	1	1	0	1	0	0	0	0
AB-p	0,00	0,00	0,24	0,04	0,32	0,22	0,84	0,28
AB-T	1	0	0	0	0	1	0	1
AB-p	0,00	0,41	0,07	0,19	0,55	0,04	0,28	0,00
AC-F	0	0	0	1	0	1	0	0
AC-p	0,69	0,65	0,13	0,00	0,06	0,00	0,85	0,74
AC-T	0	1	1	1	1	1	1	0
AC-p	0,22	0,00	0,00	0,00	0,00	0,01	0,00	0,71
BC-F	0	1	0	1	0	1	0	0
BC-p	0,11	0,05	0,45	0,00	0,22	0,00	0,77	0,34
BC-T	1	1	1	1	1	1	1	1
BC-p	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Tabulka E.7: Testování hypotéz – křížový test B, část 2

	ath	voda	ict	ect	ctv	alt	glu1	glu2
AB-F	0	0	1	0	1	1	0	0
AB-p	0,29	0,47	0,00	0,36	0,02	0,03	0,89	0,99
AB-T	1	1	1	1	1	0	0	0
AB-p	0,00	0,00	0,00	0,00	0,00	0,30	0,71	0,47
AC-F	0	0	1	1	1	0	0	0
AC-p	0,72	0,55	0,03	0,00	0,04	0,84	0,80	0,75
AC-T	0	0	0	0	0	1	1	0
AC-p	0,72	0,69	0,35	0,38	0,27	0,05	0,01	0,09
BC-F	0	0	0	1	0	0	0	0
BC-p	0,34	0,83	0,93	0,00	0,30	0,29	0,86	0,76
BC-T	1	1	1	1	1	1	1	1
BC-p	0,00	0,00	0,01	0,00	0,00	0,02	0,00	0,04

Tabulka E.8: Statistické parametry shluků – test B, část 1

	BMI	ins1	c-p1	ins2	c-p2	ins3	c-p3	tuk
$\tilde{x}(\mathbf{A1})$	34,4	53	0,60	382	2,61	241	2,39	43,00
$\tilde{x}(\mathbf{B1})$	26,8	40	0,51	312	2,52	196	2,04	31,00
$\tilde{x}(\mathbf{C1})$	36,7	120	1,26	1382	5,78	751	5,02	43,00
$\tilde{x}(\mathbf{A2})$	35,3	61	0,70	473	2,92	271	2,19	43,00
$\tilde{x}(\mathbf{B2})$	26,2	45	0,56	308	2,42	173	1,98	28,50
$\tilde{x}(\mathbf{C2})$	35,6	147	1,29	1477	5,52	1247	5,00	42,00
$\bar{x}(\mathbf{A1})$	35,9	60	0,66	418	2,57	295	2,33	42,53
$\bar{x}(\mathbf{B1})$	26,7	53	0,55	357	2,48	222	2,14	28,77
$\bar{x}(\mathbf{C1})$	38,5	124	1,25	1409	5,48	862	4,71	41,60
$\bar{x}(\mathbf{A2})$	35,8	65	0,71	516	2,96	267	2,28	43,18
$\bar{x}(\mathbf{B2})$	26,7	46	0,54	371	2,52	209	2,02	27,11
$\bar{x}(\mathbf{C2})$	37,9	180	1,53	1502	5,00	1294	5,21	42,47
$\sigma^2(\mathbf{A1})$	43,2	1105	0,08	36392	0,52	36108	0,77	46,93
$\sigma^2(\mathbf{B1})$	15,8	2863	0,11	65678	0,69	25121	0,81	34,23
$\sigma^2(\mathbf{C1})$	32,7	812	0,15	310421	1,19	241796	0,65	52,27
$\sigma^2(\mathbf{A2})$	27,3	1163	0,06	79808	0,97	26603	0,72	26,00
$\sigma^2(\mathbf{B2})$	24,5	659	0,06	64263	1,09	21768	0,63	37,12
$\sigma^2(\mathbf{C2})$	36,5	19731	0,52	349756	1,12	248116	0,90	70,64

Tabulka E.9: Statistické parametry shluků – test B, část 2

	ath	voda	ict	ect	ctv	alt	glu1	glu2
$\tilde{x}(\mathbf{A1})$	56,00	41,00	25,00	18,00	44,00	0,42	7,90	6,20
$\tilde{x}(\mathbf{B1})$	69,00	50,00	28,00	22,00	51,00	0,39	7,60	6,20
$\tilde{x}(\mathbf{C1})$	56,00	40,50	24,50	19,00	44,00	0,66	9,60	7,60
$\tilde{x}(\mathbf{A2})$	56,00	41,00	25,00	18,00	44,00	0,42	8,30	6,00
$\tilde{x}(\mathbf{B2})$	71,00	50,00	28,50	22,00	51,00	0,47	7,40	5,65
$\tilde{x}(\mathbf{C2})$	57,00	42,00	25,00	18,00	45,00	0,62	12,00	9,20
$\bar{x}(\mathbf{A1})$	56,52	41,35	24,57	18,02	43,30	0,50	8,02	6,37
$\bar{x}(\mathbf{B1})$	70,37	51,42	28,65	23,07	52,23	0,44	7,87	6,15
$\bar{x}(\mathbf{C1})$	57,40	42,00	25,60	18,50	44,80	0,72	10,12	7,36
$\bar{x}(\mathbf{A2})$	55,92	40,83	25,26	18,40	44,03	0,51	8,18	5,98
$\bar{x}(\mathbf{B2})$	71,95	51,61	30,13	22,24	52,76	0,56	7,73	5,92
$\bar{x}(\mathbf{C2})$	56,53	42,35	25,06	18,76	44,18	0,74	12,11	9,28
$\sigma^2(\mathbf{A1})$	46,59	29,59	28,11	10,83	40,08	0,10	4,43	2,24
$\sigma^2(\mathbf{B1})$	34,14	23,87	7,23	13,97	20,42	0,05	4,23	2,23
$\sigma^2(\mathbf{C1})$	52,27	20,00	7,16	1,17	10,62	0,09	3,63	2,47
$\sigma^2(\mathbf{A2})$	26,17	12,49	6,45	4,09	6,50	0,06	4,32	2,31
$\sigma^2(\mathbf{B2})$	36,70	31,60	36,06	8,51	34,40	0,10	5,70	2,02
$\sigma^2(\mathbf{C2})$	70,64	31,24	49,68	41,94	11,90	0,18	3,48	4,18

Tabulka E.10: Testování hypotéz – souhlasný test C, část 1

	BMI	fat1	fat2	fat3	ins1	c-p1	ins2	c-p2
A-F	0	0	0	0	1	0	0	0
A-p	0,14	0,15	0,47	0,26	0,03	0,21	0,96	0,27
A-T	0	0	0	0	0	0	0	0
A-p	0,19	0,89	0,90	0,56	0,42	0,26	0,54	0,88
B-F	1	0	0	0	0	0	0	0
B-p	0,05	0,66	0,59	0,71	0,76	0,22	1,00	0,65
B-T	0	0	0	0	0	0	0	0
B-p	0,39	0,28	0,21	0,43	0,65	0,11	0,57	0,16
C-F	0	0	1	0	1	0	1	1
C-p	0,96	0,29	0,03	0,68	0,03	0,07	0,02	0,03
C-T	0	1	0	0	0	0	0	1
C-p	0,97	0,01	0,59	0,53	0,66	0,66	0,14	0,04

Tabulka E.11: Testování hypotéz – souhlasný test C, část 2

	ins3	c-p3	tuk	ath	voda	bezv	ect	ctv
A-F	0	0	0	0	0	0	0	1
A-p	0,10	0,83	0,07	0,09	0,22	0,05	0,05	0,00
A-T	0	0	0	0	0	0	0	0
A-p	0,12	0,38	0,45	0,44	0,41	0,56	0,07	0,20
B-F	0	0	0	0	0	0	1	0
B-p	0,60	0,61	0,16	0,15	0,10	0,23	0,00	0,06
B-T	0	0	0	0	0	0	0	0
B-p	0,95	0,68	0,93	0,96	0,24	0,06	0,87	0,66
C-F	0	0	0	0	0	0	0	0
C-p	0,21	0,43	1,00	0,96	0,27	0,26	0,20	0,59
C-T	0	0	0	0	0	0	0	0
C-p	0,93	0,90	0,62	0,66	0,72	0,19	0,47	0,92

Tabulka E.12: Testování hypotéz – souhlasný test C, část 3

	alt	moč	ldlv	hdlc	chol	glu1	glu2
A-F	1	0	0	0	0	0	0
A-p	0,01	0,61	0,94	0,16	0,66	0,89	0,17
A-T	0	0	0	0	0	0	0
A-p	0,62	0,19	0,62	0,56	0,97	0,23	0,09
B-F	0	0	0	0	0	0	0
B-p	0,58	0,96	0,10	0,39	0,09	0,12	0,67
B-T	0	0	0	0	0	0	0
B-p	0,89	0,21	0,66	0,36	0,71	0,34	0,95
C-F	0	0	0	0	0	0	0
C-p	0,10	0,29	0,19	0,45	0,10	0,40	0,28
C-T	0	0	0	0	0	0	0
C-p	0,41	0,56	0,73	0,53	0,38	0,53	0,22

Tabulka E.13: Testování hypotéz – křížový test C, část 1

	BMI	fat1	fat2	fat3	ins1	c-p1	ins2	c-p2
AB-F	0	0	0	1	1	1	1	0
AB-p	0,09	0,62	0,16	0,02	0,03	0,00	0,00	0,06
AB-T	0	0	1	1	0	1	0	0
AB-p	0,93	0,17	0,00	0,01	0,38	0,02	0,95	0,15
AC-F	1	0	0	0	0	1	0	0
AC-p	0,03	0,26	0,79	0,45	0,11	0,02	0,25	0,83
AC-T	1	1	0	0	1	0	1	1
AC-p	0,00	0,03	0,16	0,12	0,00	0,07	0,00	0,01
BC-F	1	0	0	1	0	0	1	0
BC-p	0,00	0,58	0,20	0,05	0,42	0,21	0,00	0,05
BC-T	1	1	1	1	1	1	1	1
BC-p	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,00

Tabulka E.14: Testování hypotéz – křížový test C, část 2

	ins3	c-p3	tuk	ath	voda	bezv	ect	ctv
AB-F	0	0	1	1	1	0	1	1
AB-p	0,39	0,29	0,03	0,03	0,00	0,71	0,00	0,01
AB-T	0	0	1	1	1	1	0	1
AB-p	0,31	0,39	0,00	0,00	0,00	0,00	0,23	0,00
AC-F	1	0	0	0	1	0	0	1
AC-p	0,04	0,58	0,97	0,98	0,05	0,12	0,08	0,00
AC-T	1	0	1	1	1	1	1	1
AC-p	0,01	0,17	0,00	0,00	0,00	0,00	0,00	0,00
BC-F	1	0	1	1	0	0	1	0
BC-p	0,00	0,08	0,02	0,02	0,16	0,05	0,00	0,09
BC-T	0	0	0	0	0	1	1	0
BC-p	0,32	0,85	0,39	0,42	0,98	0,00	0,02	0,30

Tabulka E.15: Testování hypotéz – křížový test C, část C

	alt	moč	ldlv	hdlc	chol	glu1	glu2
AB-F	0	0	0	0	0	0	0
AB-p	0,78	0,96	0,27	0,16	0,13	0,74	0,75
AB-T	0	1	1	1	1	0	0
AB-p	0,11	0,01	0,03	0,00	0,00	0,22	0,09
AC-F	1	0	0	0	0	0	0
AC-p	0,00	0,15	0,19	0,14	0,84	0,52	0,72
AC-T	0	1	0	1	1	0	0
AC-p	0,29	0,00	0,22	0,00	0,01	0,71	0,86
BC-F	1	0	0	1	0	0	0
BC-p	0,00	0,15	0,94	0,01	0,14	0,80	0,99
BC-T	1	1	1	1	1	0	0
BC-p	0,00	0,00	0,00	0,00	0,00	0,35	0,10

Tabulka E.16: Statistické parametry shluků – test C, část 1

	BMI	fat1	fat2	fat3	ins1	c-p1	ins2	c-p2
$\tilde{x}(\mathbf{A1})$	35,6	43	18,50	14	62	0,70	509	2,84
$\tilde{x}(\mathbf{B1})$	35,2	61	37,35	30	68	0,96	419	3,15
$\tilde{x}(\mathbf{C1})$	28,0	29	5,55	4	37	0,50	297	2,28
$\tilde{x}(\mathbf{A2})$	35,3	41	16,80	15	82	0,78	492	2,85
$\tilde{x}(\mathbf{B2})$	35,8	45	37,00	40	79	0,75	542	2,96
$\tilde{x}(\mathbf{C2})$	27,9	49	8,90	11	47	0,65	391	2,61
$\bar{x}(\mathbf{A1})$	37,6	42	17,53	16	70	0,70	551	2,96
$\bar{x}(\mathbf{B1})$	37,4	51	40,66	31	80	0,97	557	3,37
$\bar{x}(\mathbf{C1})$	28,1	30	12,00	10	46	0,56	331	2,42
$\bar{x}(\mathbf{A2})$	35,6	43	18,07	18	79	0,78	517	2,93
$\bar{x}(\mathbf{B2})$	35,4	44	33,41	36	74	0,77	633	2,87
$\bar{x}(\mathbf{C2})$	28,1	45	10,31	12	49	0,60	402	2,89
$\sigma^2(\mathbf{A1})$	35,8	328	197,99	136	677	0,04	35208	0,47
$\sigma^2(\mathbf{B1})$	75,8	407	367,45	402	1781	0,19	185048	1,10
$\sigma^2(\mathbf{C1})$	15,7	521	222,10	185	1305	0,12	22657	0,52
$\sigma^2(\mathbf{A2})$	19,3	178	268,86	220	1756	0,07	36014	0,75
$\sigma^2(\mathbf{B2})$	30,5	334	286,82	338	1548	0,11	185350	1,36
$\sigma^2(\mathbf{C2})$	15,5	359	101,64	159	612	0,06	51262	1,13

Tabulka E.17: Statistické parametry shluků – test C, část 2

	ins3	c-p3	tuk	ath	voda	bezv	ect	ctv
$\tilde{x}(\mathbf{A1})$	405	2,84	46,00	53,00	39,00	15,00	18,00	43,00
$\tilde{x}(\mathbf{B1})$	268	1,96	31,50	67,50	48,50	20,50	19,00	49,00
$\tilde{x}(\mathbf{C1})$	243	2,40	35,00	64,00	47,00	12,00	21,00	49,00
$\tilde{x}(\mathbf{A2})$	330	2,36	44,00	55,00	40,00	14,00	18,00	44,00
$\tilde{x}(\mathbf{B2})$	247	2,00	34,00	65,00	45,00	22,00	19,00	47,00
$\tilde{x}(\mathbf{C2})$	226	2,33	35,00	64,00	47,00	12,50	21,50	48,00
$\bar{x}(\mathbf{A1})$	423	2,78	45,22	53,87	38,91	14,87	17,43	40,74
$\bar{x}(\mathbf{B1})$	341	2,48	32,65	66,35	47,90	19,85	18,90	48,10
$\bar{x}(\mathbf{C1})$	270	2,42	34,59	64,56	47,94	11,94	22,00	49,41
$\bar{x}(\mathbf{A2})$	330	2,54	44,16	54,96	39,56	14,32	18,32	43,28
$\bar{x}(\mathbf{B2})$	347	2,33	32,86	66,24	46,00	22,05	19,10	47,52
$\bar{x}(\mathbf{C2})$	267	2,39	33,91	65,15	47,59	12,76	21,56	49,50
$\sigma^2(\mathbf{A1})$	54782	0,96	30,63	30,30	8,81	14,12	3,71	83,29
$\sigma^2(\mathbf{B1})$	80364	1,54	79,40	79,40	34,41	16,56	24,94	23,67
$\sigma^2(\mathbf{C1})$	25329	0,78	30,55	30,86	19,94	7,75	7,64	12,13
$\sigma^2(\mathbf{A2})$	27348	0,87	14,31	14,87	5,26	6,23	1,64	3,21
$\sigma^2(\mathbf{B2})$	102342	1,21	41,33	40,99	16,00	9,55	3,79	10,06
$\sigma^2(\mathbf{C2})$	39497	1,03	30,51	30,37	13,52	5,22	4,86	14,68

Tabulka E.18: Statistické parametry shluků – test C, část 3

	alt	moč	ldlv	hdlc	chol	glu1	glu2
$\tilde{x}(\mathbf{A1})$	0,41	310	3,41	1,47	5,33	8,70	6,70
$\tilde{x}(\mathbf{B1})$	0,62	369	2,59	1,27	4,26	8,10	6,20
$\tilde{x}(\mathbf{C1})$	0,40	256	3,53	1,96	5,78	8,45	6,65
$\tilde{x}(\mathbf{A2})$	0,41	284	3,04	1,57	5,32	8,20	6,20
$\tilde{x}(\mathbf{B2})$	0,73	378	2,85	1,14	4,31	8,00	6,00
$\tilde{x}(\mathbf{C2})$	0,44	261	3,49	1,77	5,74	8,30	5,90
$\bar{x}(\mathbf{A1})$	0,53	312	3,28	1,53	5,31	8,76	6,96
$\bar{x}(\mathbf{B1})$	0,72	366	2,78	1,26	4,49	8,02	6,14
$\bar{x}(\mathbf{C1})$	0,43	253	3,53	1,91	5,86	8,56	6,89
$\bar{x}(\mathbf{A2})$	0,48	290	3,19	1,60	5,31	8,12	6,30
$\bar{x}(\mathbf{B2})$	0,74	391	2,68	1,18	4,39	8,75	6,17
$\bar{x}(\mathbf{C2})$	0,47	261	3,46	1,85	5,72	8,21	6,36
$\sigma^2(\mathbf{A1})$	0,16	3833	0,42	0,11	0,46	3,38	2,16
$\sigma^2(\mathbf{B1})$	0,14	3905	0,68	0,06	0,91	3,90	2,48
$\sigma^2(\mathbf{C1})$	0,03	2223	0,71	0,19	0,51	4,40	2,51
$\sigma^2(\mathbf{A2})$	0,05	3095	0,40	0,19	0,39	3,20	1,21
$\sigma^2(\mathbf{B2})$	0,11	3829	0,32	0,08	0,41	8,05	3,03
$\sigma^2(\mathbf{C2})$	0,06	3220	0,45	0,15	0,28	5,92	3,68